

Estimating the reliability model parameters through a simulation of warranty claims: How much data is needed?

Raymond Summit¹

(Received 10 January 2012; revised 12 July 2012)

Abstract

Manufactured goods are put through a quality control process that ensures an acceptable reliability level before being released to the public. However, manufacturing processes can change over time, which may affect the quality of a product. Warranty claims provide an opportunity to obtain feedback on the manufacturing process. In this study, a Weibull failure model is used to simulate the occurrence of warranty claims. The parameters of the Weibull model used to generate the data are estimated from the simulated warranty claims. The length of time required to obtain accurate parameter estimates is examined. This study is based on claims arising from the failure of a single component, but the techniques can also be used for a simple product or a subsystem of a more complex product that can be suitably

<http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/5114> gives this article, © Austral. Mathematical Soc. 2012. Published July 21, 2012. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

modelled by a Weibull distribution. Our study shows that at least two years of data are required to obtain acceptable estimates of the Weibull parameters.

Contents

1 Introduction	C452
2 Assumptions and survival analysis	C453
3 Modelling component failure	C454
4 The investigation	C454
5 Results	C456
6 Conclusions	C459
References	C463

1 Introduction

Although a new product generally undergoes reliability analysis under stress conditions before being released to the public, it is only in the field that the behaviour of a product under normal conditions is observed [1]. A warranty database contains a wealth of information that can be used to estimate and monitor field reliability of products [4, 9]. Warranty data can be used as an early warning system to monitor the manufacturing process [6]. The important issue of early detection of actionable problems that could lead to events such as product recalls has received some attention [5, 10]. This study addresses this issue by investigating how far into a warranty it is necessary

to collect data, in order to obtain an accurate estimate of the reliability of components within the product.

We simulate warranty claims over a period of up to five years. The simulated claims are based on the reliability of a single component, the reliability of which is modelled by a Weibull distribution. Using an Australian car manufacturer's warranty data, the reliability of the component is estimated. The maximum likelihood estimators of the Weibull parameters of a model of best fit are obtained in Section 3, and then used to generate simulated warranty claims in Section 4. The technique can be used on a subsystem or even on a simple product that has a Weibull reliability function. Other models can also be accommodated by using an appropriate distribution in place of the Weibull in the simulations. We are interested in establishing how far into the warranty period data should be collected before accurate estimates of the model parameters are made. This has application in the production process if a new component is introduced or a new batch commenced. The results are discussed in Section 5. We start by mentioning the need to establish the modelling assumptions.

2 Assumptions and survival analysis

We made standard assumptions in modelling the reliability of a component. These have been outlined in our other article in this volume [8].

In this article we use

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp \left[- \left(\frac{t}{\eta}\right)^{\beta} \right]$$

for the Weibull function, where β is the shape parameter and η is the scale parameter.

3 Modelling component failure

Using the manufacturer's warranty data, we model the reliability of one component. In one year, 30,138 vehicles were produced and 531 warranty claims were made that resulted in the part being replaced. We firstly obtain the Kaplan–Meier estimate of survival [3, 7] from this data, before fitting a Weibull distribution. The Kaplan–Meier survival probability to time t_i is

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{k(t_i) - d_i}{k(t_i)},$$

where $k(t_i)$ and $d(t_i)$ are the number at risk and the number of 'deaths' in time interval i , respectively. This estimator uses not only failed items, but also censored items, which are components that have not yet failed by the end of the experiment. A Weibull distribution is then fitted to the Kaplan–Meier survivals using maximum likelihood [2].

We wrote an S-Plus script based on the numerical solution provided by Xie [11]. The script uses the function `survReg` to obtain the Weibull parameter estimates of $\beta = 1.4624$ and $\eta = 17,789$. Figure 1 shows that the Weibull model provides a reasonable fit to the data.

4 The investigation

Using the values of the Weibull parameters determined in Section 3, simulated warranty claims were generated for a single component based on one production year of 30,138 vehicles. Simulations were generated for five different periods, representing data obtained over one, two, three, four and five-year warranty periods. In the simulations, the warranty is bound by time alone, in keeping with the method used to obtain the Weibull parameters from the manufacturer's data. After each simulation, a Weibull model is fitted to the data and estimates of the model's parameters made using maximum likelihood.

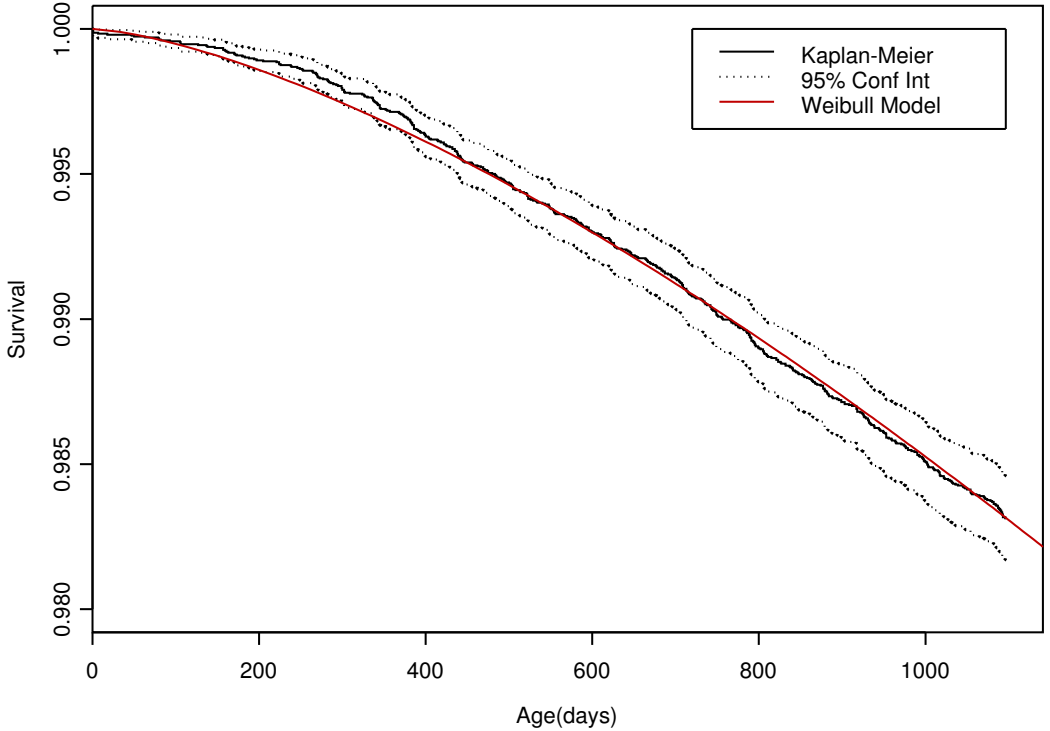


Figure 1: Kaplan–Meier and Weibull fits

In the simulations, both the vehicle’s age and the component’s age are monitored. Only components that fail in vehicles that are still within their warranty period result in a claim. The new component that replaces it has the potential to generate another claim in the database if the vehicle that it is going into is still in warranty. Thus, the simulations generate both failed and censored data, the latter arising from components that have not failed by the end of the warranty period.

From each simulation, estimates of the shape and scale parameters were made. The simulation was then repeated 1,000 times for each warranty period of one to five years. Point and interval estimates of the parameters were then

Table 1: Parameter estimates.

Warranty Period	$\hat{\beta}$	β 95% CI	$\hat{\eta}$	η 95% CI
one year	2.688	[1.24, 18.06]	27,040	[3,350, 39,910]
two years	1.474	[1.31, 1.66]	17,910	[12,350, 26,400]
three years	1.475	[1.35, 1.62]	17,640	[13,530, 22,840]
four years	1.473	[1.37, 1.58]	17,630	[14,620, 21,420]
five years	1.468 3	[1.38, 1.56]	17,750	[15,350, 20,610]

determined for each warranty length under the assumption that with this number of simulations, the parameter estimates are normally distributed.

5 Results

Using the data from the 1,000 simulations for each warranty period of one to five years, we estimate the Weibull parameters, their standard deviations and 95% confidence intervals. The results are shown in Table 1. Boxplots of the data are displayed in Figures 2 and 3, which clearly show that the one-year data have produced a wide range of estimates of both parameters. Thus, a single database comprising claims from vehicles up to one year old can produce grossly inaccurate estimates of both parameters. By two years, reasonable point and interval estimates are observed for both parameters, with fewer outliers being evident. More accurate point estimates of the parameters and narrower interval estimates can be observed as the data matures through to five years.

Let us consider why the one-year data produces unsatisfactory estimates. Examining the manufacturer's claims database, we observe that of the 531 claims made on the component in the three-year warranty, 26 of them occurred in the first year, 199 during the second year, 279 in the third and 27 beyond three years. (The manufacturer honoured the warranty beyond three years when an ongoing problem had been identified prior to the expiry of the warranty.) The

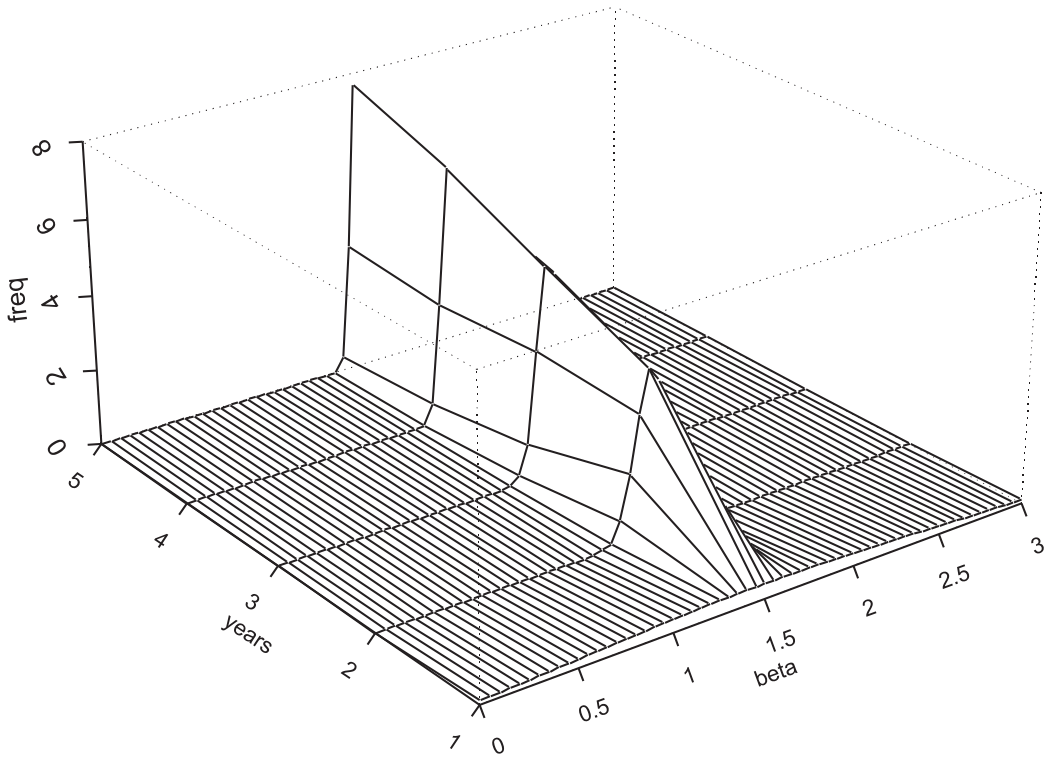


Figure 2: Shape parameter (β) boxplots.

first-year claims represent only 5% of the total claims made during the first three years. We would expect a similar proportion to occur in the simulations. Because failure time is a random variable, the differences in failure times could lead to a large variation in parameter values when fitting a Weibull model. As time increases to two or more years, the number of claims increases, enabling a better fit to the more numerous data points.

The parameter estimates from the simulations are also displayed in the 3D surface plot of Figures 4 and 5. The density plots show how the estimates hone in on their particular values as the number of years varies from one to five years. We clearly see that the interval estimates shorten over this time

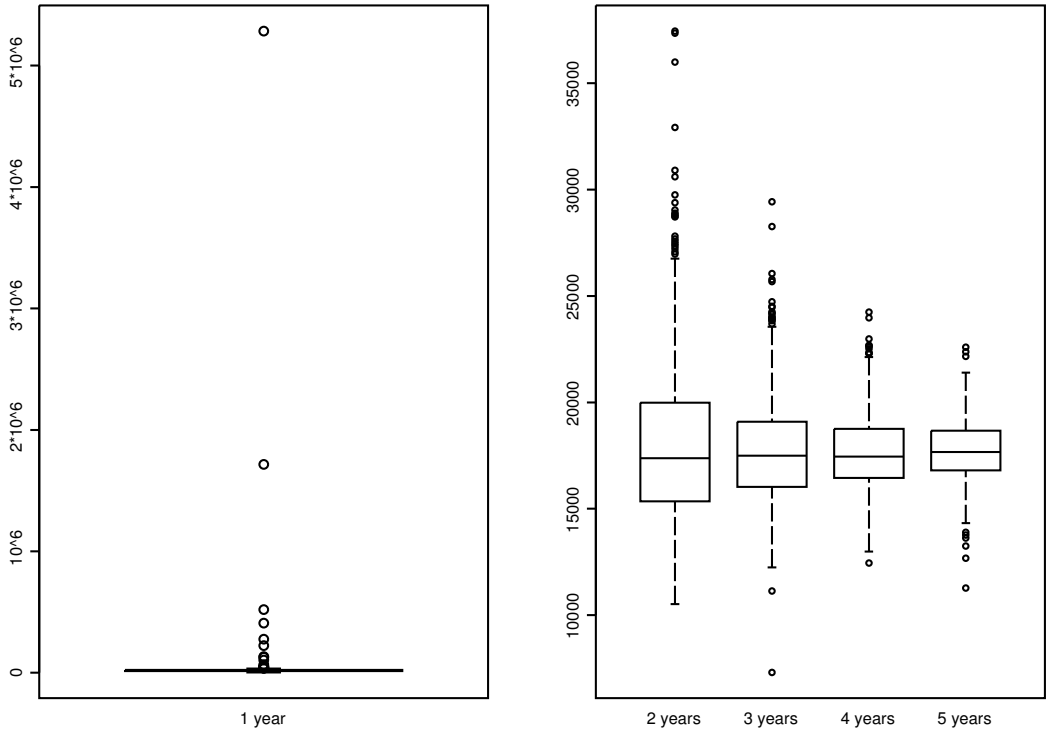


Figure 3: Scale parameter (η) boxplots.

period and that the one-year data have not provided satisfactory parameter estimates.

To test the assumption that the parameter estimates from each of the 1,000 simulations are normally distributed, quantile-quantile (QQ) plots were constructed and are displayed in Figures 6–7. The graphs clearly show that the one-year data are not normally distributed for either parameter. By two years, the β estimates are fairly normal and the η estimates are becoming more normally distributed. By three years, both parameters are reasonably normally distributed. The QQ plots for five years are omitted as they are very similar to the four-year plots, which can be taken to be normally distributed.

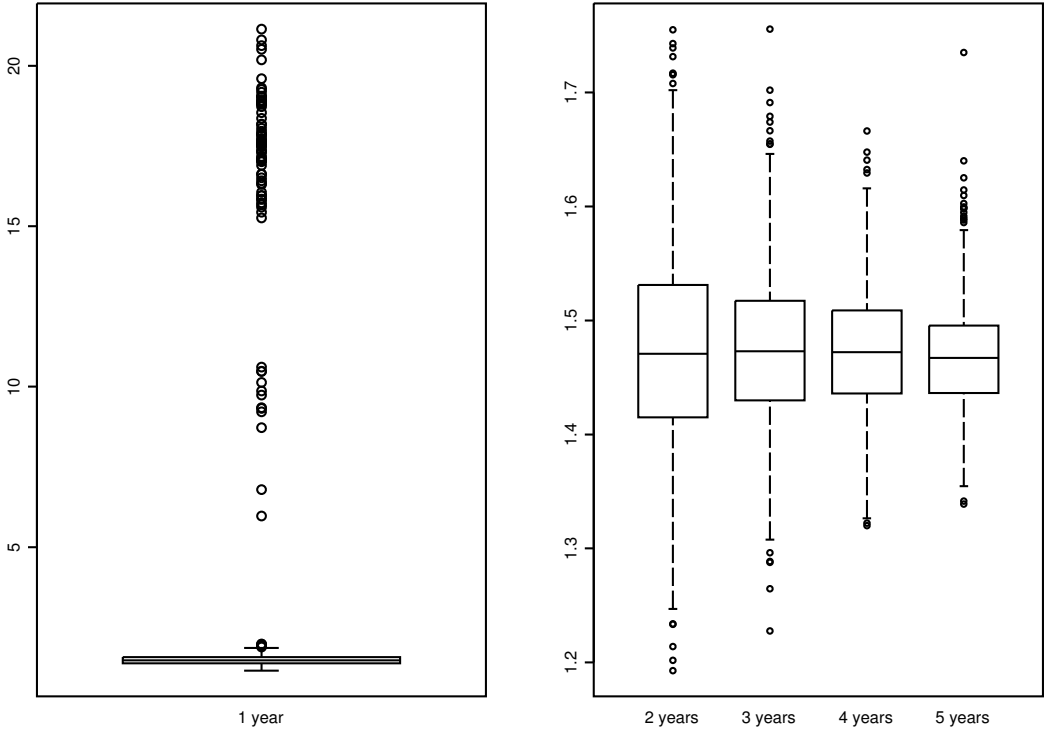


Figure 4: Shape parameter (β) surface plot.

Thus, we caution against drawing any firm conclusions from the one-year data, and possibly even from the two-year data.

6 Conclusions

We have shown that the one-year data are insufficient to produce accurate point and interval estimates of either model parameters, whereas the two-year data produced reasonably good point and interval estimates. Further improvements of point estimates of both parameters, as well as tighter interval

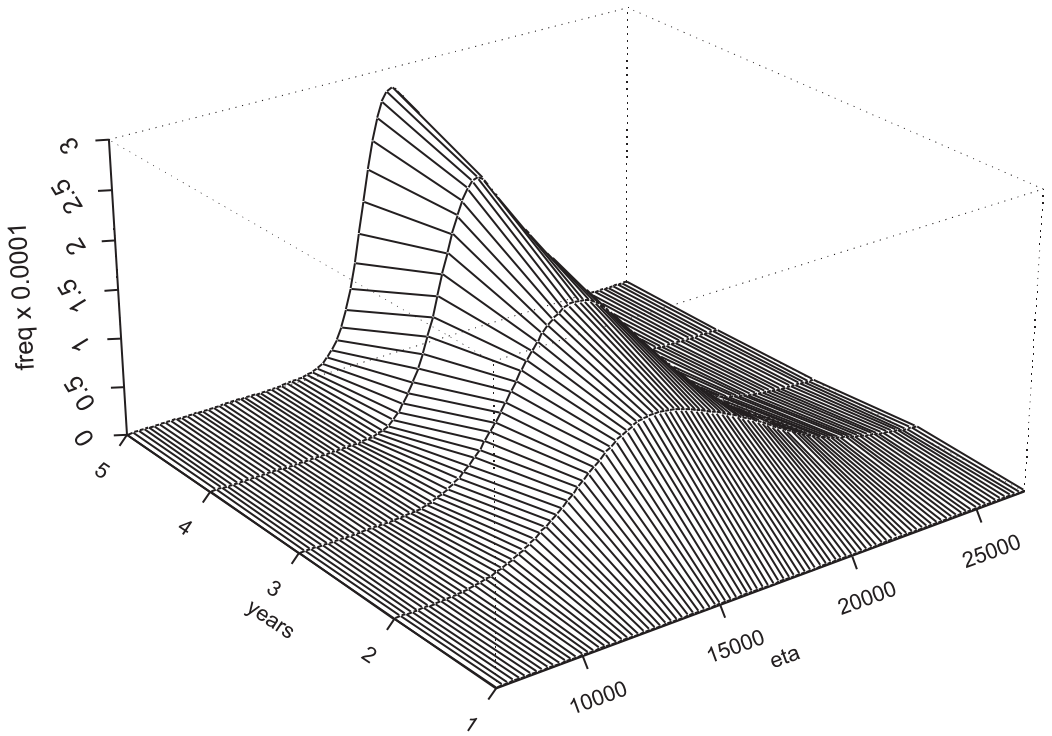


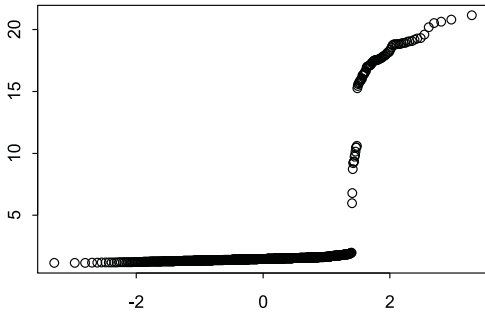
Figure 5: Scale parameter (η) surface plot.

estimates, were observed as the warranty data extended beyond two years.

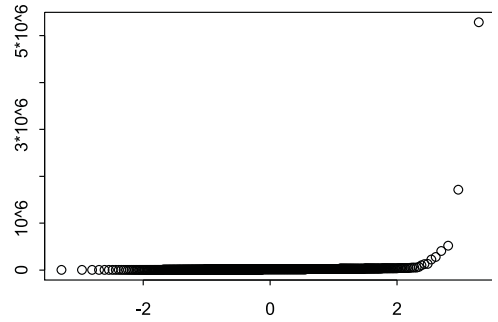
The assumption that the parameter estimates were normally distributed was shown to be inaccurate for the one-year data for both parameters and not very accurate for the η estimates, even for the two-year data. By three years, reasonably normal behaviour was observed and we conclude that accurate estimates of both parameters were obtained. Further reductions in interval widths were displayed in the four and five-year data.

We have shown that the Weibull model provided a good fit to the warranty data upon which the Weibull parameters were based in the simulations. However, other distributions could be used to model the reliability of parts.

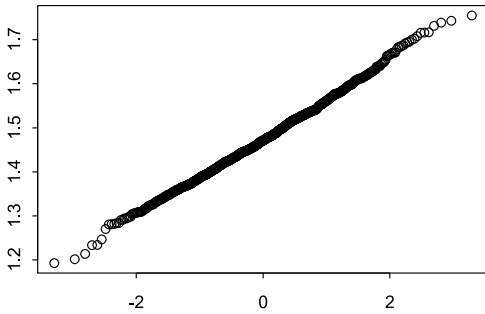
Beta normal QQ-Plot: 1 year



Eta normal QQ-Plot: 1 year



Beta normal QQ-Plot: 2 year



Eta normal QQ-Plot: 2 year

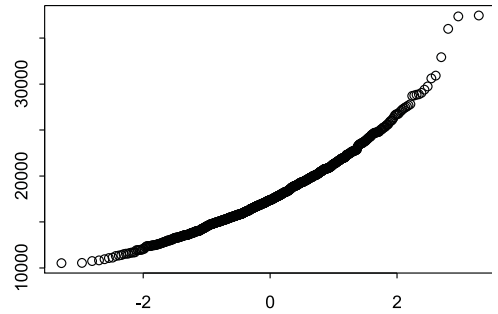
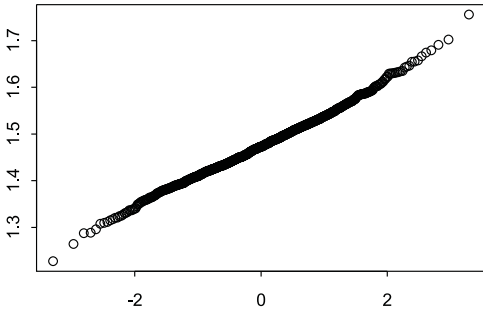


Figure 6: Sample quantiles (vertical) versus normal quantiles (horizontal).

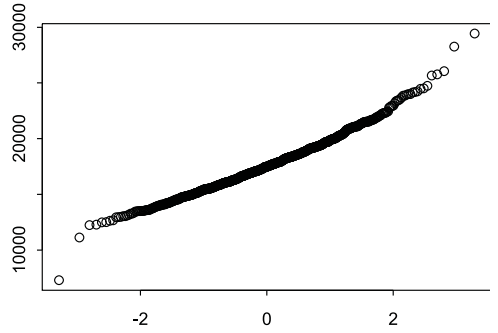
The exponential distribution may be more appropriate for components with few failures, especially if the period of observation is restricted to less than two years. We intend to repeat the experiment using the exponential model to see if stronger results are obtained.

It would be of benefit to a manufacturer to use warranty data as a means of monitoring the production process. If critical values of parameters that would result in an actionable situation are identified, warranty costs can be kept at bay.

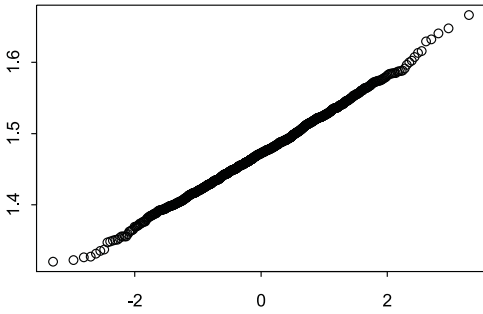
Beta normal QQ-Plot: 3 year



Eta normal QQ-Plot: 3 year



Beta normal QQ-Plot: 4 year



Eta normal QQ-Plot: 4 year

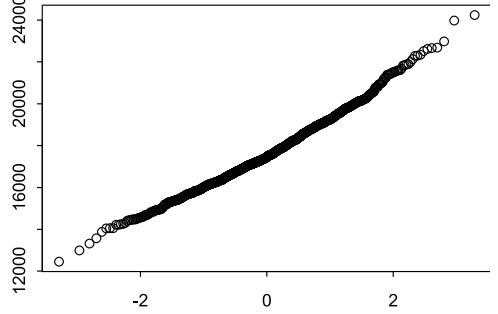


Figure 7: Sample quantiles (vertical) versus normal quantiles (horizontal).

Since warranty data does not increment in yearly intervals but is collected continuously, we have conducted a similar investigation to the one reported herein, that reflects this reality. This has led to somewhat different results to those reported in this article. We anticipate that these findings will soon be reported.

References

- [1] Campean, I. F. and Brunson, D. Reliability modelling using warranty data. *International Conference on Statistics and Analytical Methods in Automotive Engineering*, IMechE Conference Transactions 2002, 225–236. [C452](#)
- [2] Cox, D. R. and Oakes, D. *Analysis of Survival Data*. Chapman and Hall, London, 1984. [C454](#)
- [3] Lawless, J. F. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York, 1982. [C454](#)
- [4] Lawless, J. F. Statistical analysis of product warranty data. *International Statistical Reviews*, **66**, 1998, 41–60.
[doi:10.1111/j.1751-5823.1998.tb00405.x](https://doi.org/10.1111/j.1751-5823.1998.tb00405.x) [C452](#)
- [5] Lu, M. Automotive reliability prediction based on early field failure warranty data. *Quality and Reliability Engineering International*, **14**, 1998, 103–108. [doi:10.1002/\(SICI\)1099-1638\(199803/04\)14:2<103::AID-QRE147>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1099-1638(199803/04)14:2<103::AID-QRE147>3.0.CO;2-5)
[C452](#)
- [6] Menezes, M. A. J. and Quelch, J. A. Leverage your warranty program. *Sloan Management Review*, **31**, 1990, 69–80. [C452](#)
- [7] Nelson, W. *Applied Life Data Analysis*. John Wiley and Sons, New York, 1982. [C454](#)
- [8] Summit, R. Modelling component reliability using warranty data. In M. Nelson, M. Coupland, H. Sidhu, T. Hamilton, and A. J. Roberts, editors, *Proceedings of the 10th Biennial Engineering Mathematics and Applications Conference*, EMAC-2011, volume 53 of *ANZIAM Journal*, 2012, C437–C450. <http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/5115> [C453](#)

- [9] Suzuki, K. Estimation of lifetime parameters from incomplete field data. *Technometrics*, **27**, 1985, 263–271.
<http://www.jstor.org/stable/1269707> C452
- [10] Wu, H. and Meeker, W. Q. Early detection of reliability problems using information from warranty databases. *Technometrics*, **44**, 2002, 120–133.
<http://www.stat.iastate.edu/preprint/articles/2001-06.pdf>
C452
- [11] Xie, M. On the Solution of Renewal-Type Integral Equations. *Communications in Statistics*, **18**, 1989, 281–293.
[doi:10.1080/03610918908812760](https://doi.org/10.1080/03610918908812760) C454

Author address

1. **Raymond Summit**, WA School of Mines, Curtin University, Kalgoorlie, AUSTRALIA.
<mailto:r.summit@curtin.edu.au>