Comparative Distributions of Hazard Modeling Analysis

Rana Abdul Wajid Professor and Director Center for Statistics Lahore School of Economics Lahore E-mail: drrana@lse.edu.pk

M. Shuaib Khan Department of Statistics Govt. College Makhdoom Rashid Multan E-mail: skn_801@yahoo.com

Abstract

In this paper we present the comparison among the distributions used in hazard analysis. Simulation technique has been used to study the behavior of hazard distribution modules. The fundamentals of Hazard issues are discussed using failure criteria. We present the flexibility of the hazard modeling distribution that approaches to different distributions.

Key Words: Comparative Distributions, simulation, bathtub curve, instantaneous failure rate.

1. Introduction

The hazard function h(t) is the conditional probability of an item failing in the interval t to (t + dt) given that it has not failed by time t. This is also known as the instantaneous failure rate. Taking the bathtub curve, the early failure period has a decreasing hazard function as time goes by. The useful life period has a constant hazard function. The wear-out period has an increasing hazard function. The hazard function (HF) denoted by h(t), is defined as f(t)/R(t). The units for h(t) are probability of failure per unit of time, distance or cycles. The hazard rate is defined as the probability per time unit that a case that has survived to the beginning of the respective interval will fail in that interval. Specifically, it is computed as the number of failures per time units in the respective interval, divided by the average number of surviving cases at the mid-point of the interval. This is the survival time at which the cumulative survival function is equal to 0.5. Other percentiles (25th and 75th percentile) of the cumulative survival function can be computed accordingly. The 50th percentile (median) for the cumulative survival function is usually not the same as the point in time up to which 50% of the sample survived. This would only be the case if there were no censored observation prior to this time.

2. Theoretical Background

2.1 Mathematical Function of Hazard

This study is concerned with the hazard over time $t \ge 0$. The reliability function R(t) is the probability of a device not failing prior to some time t is given by

$$R(t) = 1 - F(t) = \int_{t}^{\infty} f(t) dt$$
(1)

By differentiating equation (1) it can be shown that

$$\frac{-dR(t)}{dt} = f(t)$$

The probability of failure in a given time interval between t_1 and t_2 can be expressed by the reliability function

$$R(t_{1}) - R(t_{2}) = \int_{t_{1}}^{\infty} f(t) dt - \int_{t_{2}}^{\infty} f(t) dt$$

The rate, at which failures occur in the interval (t_1, t_2) , the failure rate, $\eta(t)$ is defined as the ratio of probability that failure occurs in the interval, given that it has not occurred prior to t_1 , the start of the interval, divided by the interval length.

$$\eta(t) = \frac{R(t_1) - R(t_2)}{\Delta t R(t)}$$

where $\Delta t = t_2 - t_1$, $t = t_1$ and $t_2 = t + \Delta t$. The hazard rate h(t) is the instantaneous failure rate, is defined as the limit of the failure rate as the interval length approaches zero.

$$h(t) = \lim_{\Delta t \to 0} \left[\frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \right]$$

or
$$h(t) = \frac{f(t)}{R(t)}$$
 (2)

This is one of the fundamental relationships in hazard analysis. The density function of the time to failure f(t) and the reliability function R(t), the hazard rate function for any time t can be found. The relationship is fundamental and important because it is independent of the statistical distribution under consideration. The differential equation tells us that the hazard rate is nothing more than a measure of the change in survivor rate per unit change in time. From equation (2) a general expression is derived for hazard (failure) rate. This can also be done for the reliability function R(t).

$$\frac{dR(t)}{R(t)} = -h(t)dt$$
(3)
gives $R(t) = \exp\begin{bmatrix} t \\ -\int h(t)dt \\ 0 \end{bmatrix}$
(4)

under the conditions R(0) = 1, $\ln R(0) = 0$.

Eq.(4) is the general expression for the reliability function. If h(t) can be considered a constant failure rate η , Which is true for many cases for electronic component, Eq.(4) becomes

$$R(t) = e^{-\eta t} \tag{5}$$

Eq.(5) is used quite frequently in reliability analysis, particularly for electronic equipment. However, the reliability analyst should assume himself that the constant failure rate assumption is valid for the item being analyzed by performing goodness of fit test on the data.

2.2 Bathtub curve

In practice, the patterns of failures over time are often classified into infant mortality, useful life and wear-out. These patterns can be recognized in mathematics by a combination of decreasing, constant, and increasing hazard function. The three patterns combine to produce the well-known bathtub curve shown in Fig. 1, Willey (1992). With a decreasing hazard function (i.e. a new item in this life period has a larger probability of failing than an old item), the left- hand section of the curve is called the infant mortality period



Life (units can be time. distance, cycles, etc.)

Fig. 1 Bathtub curve

During this period the poor quality components may be weeded out by a 'burn-in' process, which is often used in manufacturing, pre-delivery testing or is experienced very soon after an item comes into service. With a relatively even segment, the middle section of the curve is called the *useful life period*. In this zone, the hazard function seems to be stable. During this period the component has a constant hazard function, i.e. a new item in this life period has an equal probability of failing as an old item. Most of the components and systems, which fail in service, exhibit a hazard function, which takes a constant value, as shown by this section of the curve. The right-hand section of the curve, called the wearout period, has an increasing hazard function, i.e. a new item in this life period the failure is generally caused by factors such as fatigue, degradation, wear and the like.

3. Distributions of Hazard Analysis

3.1 Analysis Output and Results

In Weibull distribution when $\beta = 1$, the distribution is the same as the exponential distribution for a constant hazard function and $h_w(t) = \frac{1}{n}$ so the exponential distribution is a special case of the Weibull distribution and the Weibull distribution can be treated as a generalization of the exponential distribution. When $\beta < 1$, the hazard function is continually decreasing which represents early failures. When $\beta > 1$, the hazard function is continually increasing which represents wear-out failures- In particular, when $\beta = 2$, it is known as the Rayleigh distribution. When $\beta = 3.4$ is the shape of the PDF is similar to the Normal PDF. These cases are called pseudo-symmetrical cases by Gumbel (1958). So the Weibull is a very flexible distribution. The two parameter generalized exponential (GE) distribution has been introduced and studied guite extensively by the authors (Gupta and Kundu, 1999, 2001). The different shapes of generalized exponential (GE) distribution are guite similar to Weibull density functions. When shape parameter $\beta = 1$, it coincides with the exponential distribution. The hazard function of a generalized exponential (GE) distribution can be increasing, decreasing or constant depending on the shape parameter similarly as a Weibull distribution. Therefore, generalized exponential (GE) distribution and Weibull distributions are both generalization of an exponential distribution in different ways. If it is known or apparent from the histogram that the data are coming from a right tailed distribution, then a generalized exponential (GE) distribution can be used quite effectively. It is observed that in many situations generalized exponential (GE) distribution provide better fit than a Weibull distribution. Therefore to analyze a skewed lifetime data an experimenter might wish to chose one of the two models. From the view of statistics, if the hazard functions of the underlying distribution approximates to any part of the bathtub curve (decreasing, constant or increasing hazard function) then this distribution might be applicable as a time-to-failure model. So certain distributions, which are often used in reliability and hazard, are investigated here.

Such distributions include the Uniform, Normal, Exponential, Rayleigh, Weibull, Erlang, Gamma, generalized exponential (GE) distribution, Loglogistic, Lognormal and others. The Uniform distribution has a rapidly increasing hazard function. The Normal distribution has a gradually increasing hazard function. The Exponential distribution has a constant hazard function. The Rayleigh distribution has an increasing hazard function and is the special case of the Weibull distribution with shape parameter, $\beta = 2$. The Erlang distribution has a very gradually increasing hazard function and is the special case of the Gamma distribution. The Weibull and Gamma distributions have decreasing hazard functions for 0 < shape parameters < 1, and constant hazard functions for shape parameters = 1, and increasing hazard functions for shape parameters > 1. Note that the Exponential distribution is a special case of the Weibull, Gamma and generalized exponential (GE) distributions. The Loglogistic distribution has a decreasing hazard function when the shape parameter < 1 and has one mode when the shape parameter > 1 Cox. and Oakes, (1984). The Lognormal distribution has a plausible increasing hazard function when the shape parameter $\beta \ge 4$, but a plausible decreasing hazard function when $\beta \le 0.4$. When $\beta \approx 0.8$ to 2.5, the hazard function is near constant over most of the distribution. Hence the Uniform, Normal, Exponential, Rayleigh, Weibull, Erlang, Gamma, Loglogistic, Lognormal distributions all have appropriate hazard functions for modeling different phases in the life of a component. The hazard functions of the following distributions are shown in fig.(2) where the legends, from the top to the bottom, are Uniform(0, 4), Normal(0, 1), Exponential (1), Rayleigh (1), Weibull (1, 0.5), Weibull(1, 3), Earling(1), Loglogistic (1, 0.5), Loglogistic (1, 3), Lognormal (1, 4), Lognormal (1, 0.8) and Lognormal (1, 0.4) respectively. The value(s) in parenthesis is (are) the parameter(s) which are used to describe for each distribution.





Pak. j. stat. oper. res. Vol.II No.2 2006 pp127-134

Rana Abdul Wajid, M. Shuaib Khan

From these statistical models, the Weibull and Gamma distributions cover all the phase of component life and they can represent all the time-to-failure data. However, the Weibull distribution is the most widely distribution used in industry. This is an important advantage of the life time distributions for modeling reliability and hazard function. The Weibull distribution is the most widely distribution used in industry. Two main reasons for the popularity of the Weibull distribution are that it has simple expressions and closed forms to model the probability density function (PDF), reliability function, cumulative distribution function (CDF) and hazard function. The Gamma distribution is not as popular as the Weibull distribution partly because its (CDF and hazard function are not expressible in a simple closed form and further it is very complicated for the case of dealing with incomplete data. Lawless, (1982) describe the Gamma distribution is therefore harder to handle than the Weibull distribution. The Lognormal distribution cannot represent lifetimes exactly; nevertheless it is often used satisfactorily to match shorter values of time-to-failure Lawless, (1982). The Lognormal distribution enables plausible and acceptably accurate conclusions to be drawn about some types of failure Lawless, (1982). The relationships between various distributions are shown in Fig. 3 where the direction of each arrow represents going from the general to a special case.



Fig. 3 The relationship among various distributions

There are many publications that individually cover these distributions. However; few papers compare the differences among these distributions. All distributions, interestingly, can have frequency curves which are skewed (the Weibull can be positively or negatively skewed but the Lognormal can only be positively skewed). The generalized exponential (GE) distribution, Gamma and Weibull distributions are all generalization of an exponential distribution in different ways. If it is known or apparent from the histogram that the data are coming from a right tailed distribution, then a generalized exponential (GE) distribution can be used quite effectively. It is observed that in many situations generalized exponential (GE) distributions.

4. Conclusions

In this paper we have discussed the problem of Comparative Distributions of Hazard Modeling between overlapping families of life time distributions functions. The life time distributions are extensively used in reliability and hazard life testing. The life time distributions are fitted the life data very well. From the above comparison of these Comparative Distributions of Hazard Modeling Analysis we conclude that the Weibull, Erlang, Gamma, generalized exponential (GE) distributions are flexible distributions that approaches to the different distributions of hazard modeling. The Lognormal distribution cannot represent lifetimes exactly. It is used satisfactorily to match shorter values of time-to-failure data.

References

- 1. Abernathy, R. B. (2004). The New Weibull Handbook. Dept at Houston, Texas 77252-2608, USA.
- 2. Adams, J. D. (1962). Failure Time Distribution Estimation, Semiconductor Reliability. Vol. 2, (41-52).
- 3. Cox, D. R. and Oakes, D. (1984). Analysis of Survival Data, Chapman & Hall.
- 4. Dey, K. A. (1983). Practical Statistical Analysis for the Reliability Engineer, IT Research Institute, Rome Air Development Centre, Griffins AFB, NY 13441, USA.
- 5. Goldthwaite, L. R. (1961). Failure Rate Study for the Lognormal Lifetime Model, Proceedings, of the 7th National Symposium on Reliability and Quality Control in Electronics. (208-213).
- 6. Gumbel, E.J. (1958), Statistics of Extremes, Columbia University. Press, New York Gupta, R.D., Kundu, D., (1999). Generalized exponential distributions. Austral. N.Z.J. Statist . 41(2), (173-188).
- 7. Gupta, R.D. and Kundu, D. (2003). Discrimination between the Weibull and Generalized exponential distributions. Computational Statistics& data analysis 43 (179-196).
- 8. Gupta, R.D., Kundu, D., (2001a). Generalized exponential distributions: an alternative to Gamma and Weibull distributions. Biometrical J. 43 (1), (117-130).

- 9. Gupta, R.D., Kundu, D., (2001b). Generalized exponential distributions: different method of estimations. J. Statist. Compute Simulations 69 (4), (315-338).
- 10. Johnson, N. L. and Kotz, S. (1970). Continuous Distributions, J. Wiley. New York, USA. [ISBN: 0-471-44626-2].
- 11. Lawless, J. F. (1982). Statistical Models and Methods of life time data, J WILEY, New York, USA.
- 12. Nelson, W. (1983). Haw to Analyze Reliability Data, The ASQC Basic References in Quality Control: Statistical Techniques. Vol. 6.
- 13. Richard E. Barlow, Albert W. Marshall, Frank Proschan, Properties of Probability Distributions with Monotone Hazard Rate ,*Annals of Mathematical Statistics*, Vol. 34, No. 2 (Jun., 1963), pp. 375-389
- 14. Willey, P.C.T. (1992). Introduction to Weibull Analysis. Lecture Notes of the Department ME & OM, University of Nottingham, UK.