International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 16, 2017, pp 103-114

An Introduction to Competing Risk Model for Analysis of Reliability Data

Md. Rezaul Karim

Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh
Email: mrezakarim@yahoo.com.

[Received May 13, 2017; Accepted October 25, 2017]

Abstract

The complexity of products has been increasing with technological advances. As a result, a product may fail in different ways or causes, which are commonly known as failure modes. Competing risk model is appropriate for modeling component failures with more than one failure modes. In this paper the competing risk model is applied for analysing product reliability data with multiple failure modes. Maximum likelihood estimation method is used to estimate the parameters and various characteristics of the model and to assess and predict the reliability of the product.

Keywords: Competing risk model; Failure mode; Reliability.

AMS Classification: 62N05, 93A30.

1. Introduction

According to (ISO 8402, 1994), a product can be tangible (e.g. assemblies or processed materials) or intangible (e.g., knowledge or concepts), or a combination thereof. A product can be either intended (e.g., offering to customers) or unintended (e.g., pollutant or unwanted effects). This paper considers tangible products, specifically manufactured goods.

The complexity of products has been increasing with technological advances. As a result, a product must be viewed as a system consisting of many elements and capable of decomposition into a hierarchy of levels, with the system at the top level and parts at the lowest level. There are many ways of describing this hierarchy. One such is the nine-level description shown in Table 1, based on a hierarchy given in Blischke and Murthy (2000) and Blischke, Karim and Murthy (2011).

Table 1: Multilevel decomposition of a product

Level	Characterization
0	System
1	Sub-system
3	Assembly
4	Sub-assembly
5	Module
6	Sub-module
7	Component
8	Part

The number of levels needed to describe a product from the system level down to the part level depends on the complexity of the product. Many units, systems, subsystems, or components have more than one cause of failure. For example, (i) A capacitor can fail open or as a short, (ii) Any of many solder joints in a circuit board can fail, (iii) A semi conductor device can fail at a junction or at a lead, (iv) A device can fail because a manufacturing defect (infant mortality) or because of mechanical wear out, (v) For an automobile tire, tread can wear out or the tire may suffer a puncture.The Competing risk model is appropriate for modeling component failures with more than one mode of failure. A failure mode is a description of a fault. It is sometimes referred to as fault mode. Failure modes are identified by studying the (performance) function. Assume a (replaceable) component or unit has K different ways it can fail. These are called failure modes and underlying each failure mode is a failure mechanism. Each mode is like a component in a series-system.

Improving reliability of product is an important part of the larger overall picture of improving product quality. Therefore, in recent years many manufacturers have collected and analyzed field failure data to enhance the quality and reliability of their products and to improve customer satisfaction. This paper applies the competing risk model to analyze product failure data and to assess and predict the reliability of the product.

The remainder of the article is organized as follows: Section 2 describes competing risk model formulation. Section 3 applies the competing risk model for analyzing a set of product failure data. Section 4 concludes the paper with additional implementation issues for further research.

2. Competing risk model formulation

The cumulative density function (cdf) of the lifetime variable T of a general K-fold competing risk model is given by

$$F(t) \equiv F(t;\theta) = 1 - \prod_{k=1}^{K} [1 - F_k(t;\theta_k)] \tag{1}$$

where $F_k(t) \equiv F_k(t;\theta_k)$ are the cdf's of the K sub-populations with parameters $\theta_k, 1 \le k \le K$. Here $\theta \equiv \{\theta_k, 1 \le k \le K\}$ and we assume that $K \ge 2$.

This is called a "competing risk model" because it is applicable when an item (component or module) may fail by any one of K failure modes, i.e., it can fail due to any one of the K mutually exclusive causes in a set $\{C_1, C_2, \ldots, C_K\}$ (Blischke, et al., 2011). The competing risk model has also been called the *compound model*, *series system model*, and *multi-risk model* in the reliability literature. Let T_k be a positive-valued continuous random variable denoting the time to failure if the item is exposed only to cause $C_k, 1 \le k \le K$. If the item is exposed to all K causes at the same time and the failure causes do not affect the probability of failure by any other mode, then the time to failure is the minimum of these K lifetimes, i.e., $T = \min\{T_1, T_2, \ldots, T_K\}$, which is also positive-valued, continuous random variable.

Let $R(t)$, $h(t)$, and $H(t)$ denote the reliability, hazard, and cumulative hazard functions associated with $F(t)$, respectively, and let $R_k(t)$, $h_k(t)$, and $H_k(t)$ be the reliability function, hazard function and cumulative hazard function associated with the cdf of the k^{th} failure mode, $F_k(t)$, respectively. It can be easily shown that

$$R(t) = \prod_{k=1}^{K} R_k(t) \tag{2}$$

$$H(t) = \sum_{k=1}^{K} H_k(t) \tag{3}$$

and

$$h(t)=\sum_{k=1}^{K} h_k(t) \tag{4}$$

Note that for independent failure modes, the reliability function for the item is the product of the reliability functions for individual failure modes (2) and the hazard function for the item is the sum of the hazard functions (4). The density function of *T* is given by

$$f(t)=\sum_{k=1}^{K}\left\{\prod_{\substack{j=1\\ j\neq k}}^{K}[1-F_j(t)]\right\} f_k(t),\ \ t\geq 0 \tag{5}$$

which may be rewritten as

$$f(t)=R(t)\left\{\sum_{k=1}^{K}\left[\frac{f_k(t)}{R_k(t)}\right]\right\},\ \ t\geq 0\,. \tag{6}$$

Suppose that a component has *K* failure modes and that the failure modes are statistically independent. We look first at the general case in which the failure modes of some of the failed items are known and those of the remaining are unknown. In addition, we assume that it is not possible to determine the failure modes (or causes of failure) for the censored (non-failed) items.

Two special cases of interest (Blischke, et al., 2011) are as follows:

Case (i): The failure modes are known for all failed items.

Case (ii): The failure modes are unknown for all failed items.

Let n_1 be the number of failed units and n_2 the number of censored units. For the failed units, the post-mortem outcome is uncertain, that is, the failure modes for some units may not be known. Out of the n_1 failed items, let n_{1k} denote the number of items with failure mode $k, 1\leq k\leq K$, and $n_{10}=n_1-\sum_{k=1}^{K} n_{1k}$ the number of failures for which there is no information regarding the failure mode. Let t_{kj} denote the lifetime of the j^{th} item failing from failure mode *k*, and $\tilde{t}_i$ the i^{th} censoring time. *Note:* For Case (i), $n_{10}=0$, and for Case (ii) $n_{10}=n_1$.

For the general case, n_{1k} units out of n failed due to failure mode *k*, with failure times $\{t_{k1},t_{k2},\ldots,t_{kn_{1k}}\}$, and there are n_{10} units with failure times $\{t'_1,t'_2,\ldots,t'_{n_{10}}\}$ for which there is no information regarding the failure mode. In addition, there are $n_2 = n - \sum_{k=1}^{K} n_{1k} - n_{10}$ censored units, with censoring times $\{\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{n_2}\}$. Thelikelihood function in the general case is given by

$$L(\theta) = \prod_{k=1}^{K} \left[\prod_{j=1}^{n_{1k}} f_k(t_{kj}) \prod_{\substack{l=1 \\ l \neq k}}^{K} R_l(t_{kj}) \right] \times \prod_{k=1}^{K} \left[\prod_{j=1}^{n_{10}} f_k(t'_j) \prod_{\substack{l=1 \\ l \neq k}}^{K} R_l(t'_j) \right] \times \prod_{i=1}^{n_2} \prod_{k=1}^{K} R_k(\tilde{t}_i) . \quad (7)$$

The MLEs of the parameters are obtained by maximizing the likelihood function (7). For most distributions the ML estimation method requires numerical maximization because of the lack of closed form solutions for the estimators.

The results for the two special cases are as follows:

Case (i): The expression for the likelihood function is given by (7) with the second term equal to unity, so that

$$L_1(\theta) = \prod_{k=1}^{K} \left[\prod_{j=1}^{n_{1k}} f_k(t_{kj}) \prod_{\substack{l=1 \\ l \neq k}}^{K} R_l(t_{kj}) \right] \times \prod_{i=1}^{n_2} \prod_{k=1}^{K} R_k(\tilde{t}_i) . \quad (8)$$

Case (ii): The expression for the likelihood function is given by (7) with the first term of equal to unity,

$$L_2(\theta) = \prod_{k=1}^{K} \left[\prod_{j=1}^{n_{10}} f_k(t'_j) \prod_{\substack{l=1 \\ l \neq k}}^{K} R_l(t'_j) \right] \times \prod_{i=1}^{n_2} \prod_{k=1}^{K} R_k(\tilde{t}_i) . \quad (9)$$

The cause-specific (or failure mode-specific) hazard function for cause *k* can be written as

$$\tilde{h}_k(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t, C = k \mid T \geq t)}{\Delta t} = \frac{f_k(t)}{R(t)}, \quad (10)$$

where $f_k(t)$ is the cause-specific pdf at time *t* that represents the unconditional probability of failure of an unit at time *t* from cause *k*, and $R(t)$ is the overall reliability function representing the probability of surviving from all causes up to time *t*. Relationship (10) implies that

$$f_k(t) = \tilde{h}_k(t) R(t) . \quad (11)$$

Using (11) and (2), we can rewrite the likelihood functions (8) and (9), respectively as

$$L_1(\theta) = \prod_{k=1}^{K}\left[\prod_{j=1}^{n_{1k}} \tilde{h}_k(t_{kj})R(t_{kj})\right] \times \prod_{i=1}^{n_2} R(\tilde{t}_i) \tag{12}$$

and

$$L_2(\theta) = \prod_{k=1}^{K}\left[\prod_{j=1}^{n_{10}} \tilde{h}_k(t'_j)R(t'_j)\right] \times \prod_{i=1}^{n_2} R(\tilde{t}_i)\,. \tag{13}$$

The MLEs of the parameters of the models are obtained by maximizing (8) or (12) for Case (i) and (9) or (13) for Case (ii). More details on the formulations and applications of competing risk models can be found in Murthy, et al. (2004) and Blischke, et al. (2011).

3. Examples

This section describes the following two examples.

3.1 Exponential distribution

Suppose that K = 2, and the lifetimes of failure modes 1 and 2 independently follow exponential distributions with parameters λ_1 and λ_2, respectively. Time to failure is modeled by (1). We consider Case (i). The data consist of n units, with n_{11} units failing due to failure mode 1 with failure times $\{t_{11}, t_{12}, \ldots, t_{1n_{11}}\}$, n_{12} units failing due to failure mode 2 with failure times $\{t_{21}, t_{22}, \ldots, t_{2n_{12}}\}$, and $n_2 = n - n_{11} - n_{12}$ units censored, with censoring times $\{\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{n_2}\}$.

In this case, from (2), we have $R(\tilde{t}) = R_1(\tilde{t})R_2(\tilde{t}) = \exp\left(-(\lambda_1 + \lambda_2)\tilde{t}\right)$ and using this in (12), the log-likelihood function becomes

$$\log L = n_{11}\log(\lambda_1) - (\lambda_1+\lambda_2)\sum_{j=1}^{n_{11}} t_{1j} + n_{12}\log(\lambda_2) - (\lambda_1+\lambda_2)\sum_{j=1}^{n_{12}} t_{2j} - (\lambda_1+\lambda_2)\sum_{i=1}^{n_2}\tilde{t}_i \tag{14}$$

From this, the ML estimators of λ_1 and λ_2 are found to be

$$\hat{\lambda}_i = \frac{n_{1i}}{\sum_{j=1}^{n_{1j}} t_{1j} + \sum_{j=1}^{n_{2j}} t_{2j} + \sum_{i=1}^{n_2}\tilde{t}_i}\,,\ i = 1, 2 \tag{15}$$

It follows from (2) that the maximum likelihood estimate of the reliability function of the component is

$$\hat{R}(t) = \exp\left(-(\hat{\lambda}_1 + \hat{\lambda}_2)t\right), t \geq 0 \tag{16}$$

We consider an electronic component for which lifetimes follow an exponential distribution. The component exhibits a new mode of failure due to mounting problems. If incorrectly mounted, it can fail earlier, and this is also modeled by an exponential distribution. The parameters of the exponential distributions for failure modes 1 and 2 are λ_1 = 0.0006 and λ_2 = 0.0004 per day. From (16), the maximum likelihood estimate of the reliability function of the component is $\hat{R}(t)$ = exp(-(0.0006+0.0004)t) = exp(-0.001t),$t \geq 0$.

Figure 1 displays a comparison of the estimated reliability functions for failure mode 1, failure mode 2 and combined failure modes 1 and 2 for $0 \leq t \leq 10000$ days.

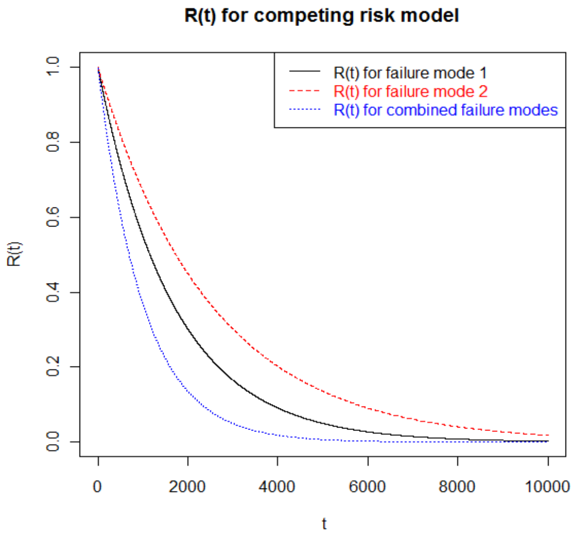

Figure 1: Comparison of ML estimates of reliability functions for competing risk model

This figure can be used to assess reliability of the component for given days. For example, the figure indicates the reliabilities of the component at age 2000 days are 0.30 for failure mode 1, 0.45 for failure mode 2 and 0.14 for the combined failure modes. Based on (16), the estimated MTTF of the component is found to be $\hat{\mu} = \int_0^\infty \hat{R}(t)dt$ = $1/(\hat{\lambda}_1 + \hat{\lambda}_2)$ = 1000 days. More details on this example is given in Blischke, et al. (2011).

3.2. Example with Device-G data

Failure times and running times for a sample of devices from a field tracking study of a larger system are given in Meeker and Escobar (1998). 30 units were installed in typical service environments. Cause of failure information was determined for each unit that failed (lifetime in thousand cycles of use). Mode S failures were caused by failures on an electronic component due to electrical surge. These failures predominated early in life. Mode W failures, caused by normal product wear, began to appear after 100 thousand cycles of use. The purposes of the analyses are:

- Analyze the failure modes separately to investigate the effects of failure modes.
- How to improve product reliability – if one failure mode can be eliminated.
- Compare lifetime (with respect to the MLEs of parameters, MTTF, B10 life, median life, etc.) of the product with failure modes (competing risk model) and ignoring failure mode information.

When the failure modes S and W act independently, one can:

- Analyze the mode S failures only: In this case mode W failures are treated as right censored observations. This is the estimate of the failure-time distribution if mode W could be completely eliminated.
- Analysis of the mode W failures only: In this case mode S failures are treated as right censored observations. This is the estimate of the failure-time distribution if mode S could be completely eliminated.
- A combined analysis use the competing risk model assuming independence between mode S and mode W.

Out of 30 units, there are 8 censored units at censoring time 300 kilocycles. A preliminary analysis of failure modes are given in Table 2. It is an examination of failure mode frequency or relative frequency data to determine the most important failure modes that contribute to quality problems and to which quality improvement efforts should be directed.

Table 2: Frequencies and average lifetimes for failure modes S and W

Failure Mode	Frequency	Average Life (Failure only)
S	15	86.1
W	7	231.3

Table 2 indicates that failure mode S has considerably higher frequency and lower average lifetime (based on failure data only). Therefore, we may conclude that efforts should be concentrated on failure mode S to eliminate it or to reduce the risks associated with this failure mode. Figure 2 represents the Weibull probability plots for individual failure modes S and W with the MLEs of shape and scale parameters. This figure suggests that the Weibull distribution provides a good fit to both failure modes.

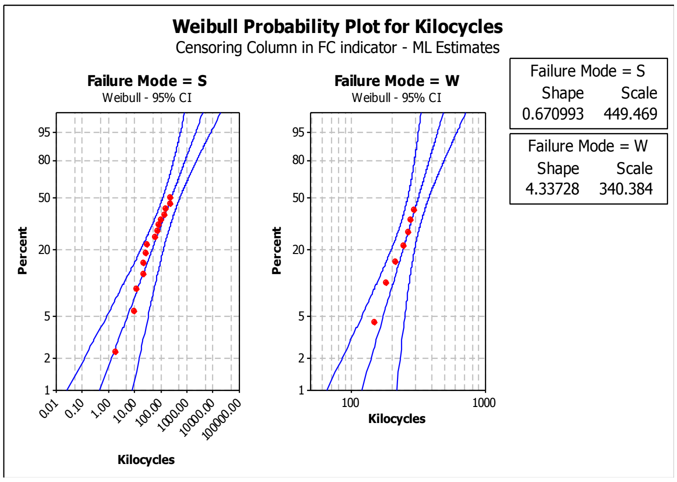

Figure 2: The Weibull probability plots for individual failure modes S and W

The maximum likelihood estimates of Weibull parameters with MTTFs for failure modes S and W are displayed in Table 3 and Table 4, respectively.

Table 3: Maximum likelihood estimates of Weibull parameters for failure mode S

Parameters and MTTF	Estimate	Standard Error	95.0% Normal CI	
			Lower	Upper
Shape	0.670993	0.157777	0.423221	1.06382
Scale	449.469	191.944	194.625	1038.01
Mean(MTTF)	593.462	342.422	191.539	1838.77

Table 4: Maximum likelihood estimates of Weibull parameters for failure mode W

Parameters and MTTF	Estimate	Standard Error	95.0% Normal CI	
			Lower	Upper
Shape	4.33728	1.45059	2.25183	8.35411
Scale	340.384	36.139	276.437	419.124
Mean(MTTF)	309.963	29.8906	256.582	374.45

Tables 3 and 4 indicate that for the failure mode W, the MLEs of shape parameter is much larger and the MTTF is smaller than that of the failure mode S. The estimates of MTTFs of Tables 3 and 4 suggest a contradiction to the conclusion taken based on the conditional average lifetimes given in Table 2 and thus it requires more investigation.

Figure 3 represents the Weibull probability plots for individual failure modes in the same scale. It suggests that the mode S failures predominated early in life whereas the mode W failures caused by normal product wear and began to appear after 100 thousand cycles of use.

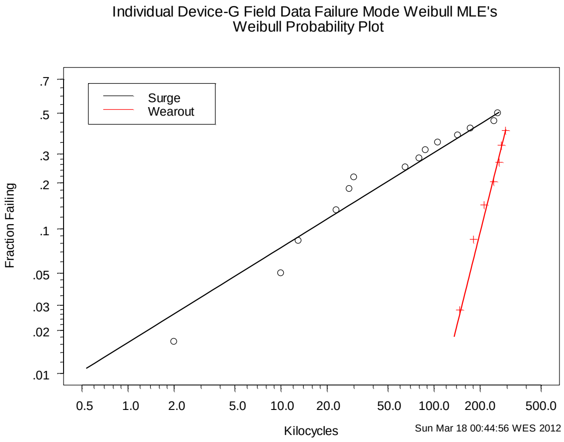

Figure 3: Weibull probability plots for individual failure modes in the same scale

Figure 4 shows the Weibull probability plot for competing risk model. This figure diverges rapidly after 200 thousand cycles.

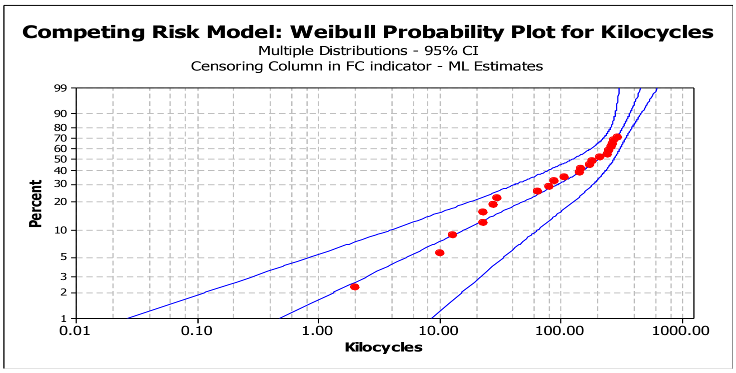

Figure 4: Weibull probability plot for competing risk model

The Weibull probability plot (ignoring failure mode information) is shown in Figure 5. Weibull analysis ignoring the failure mode information (Figure 5) shows evidence of a change in the slope of the plotted points, indicating a gradual shift from one failure mode to another.

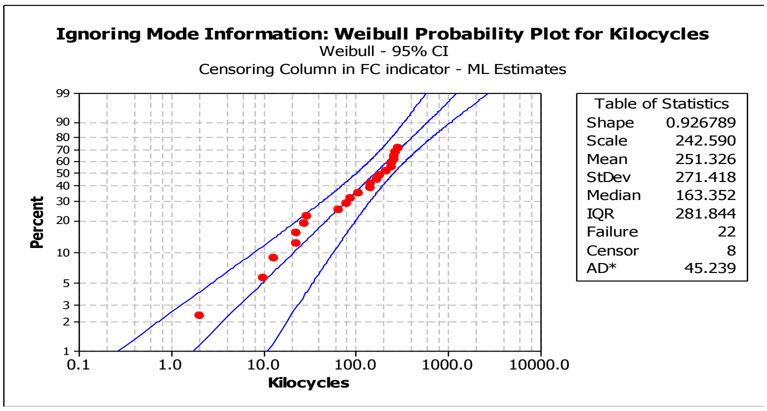

Figure 5: Weibull probability plot (ignoring failure mode information)

Maximum likelihood estimates of percentiles for both competing risk model and ignoring failure mode information are given in Table 5. From Table 5, we may conclude that, 10% of the total components fail at 15.71 kilocyclesunder competing risk model and at 21.4kilocycles under ignoring failure mode information. 50% of the total components fail at 203.06 kilocycles for competing risk model and at 163.35kilocycles for without failure mode information. Hence we may say thatignoring failure mode information over estimates the B10 life and B90 life and under estimates median life compared with the competing risk model. More on the analysis of this data set can be found in Meeker and Escobar (1998).

Table 5: MLEs of percentiles for competing risk model and ignoring failure mode information

Percentile	Competing Risk Model			Ignoring Mode Information		
	Estimate	95% L-CI	95% U-CI	Estimate	95% L-CI	95% U-CI
5	5.37	0.85	33.78	9.84	2.81	34.44
10	15.71	3.86	63.63	21.4	7.97	57.43
50	203.06	124.25	273.72	163.35	102.47	260.4
90	369.4	280.7	455.89	596.63	334.03	1065.67

4. Conclusion

- The failure mode-wise frequencies and conditional mean lifetimes can be misleading to determine the most important failure modes that contribute to quality problems and to which quality improvement efforts should be directed.
- The failure mode or failure cause wise model with competing risk is better than combined model for assessing and predicting reliability of the product.
- This article analysed the failure data based on Case (i), where the failure modes are known for all failed items. If, the failure modes are unknown for all failed items, application of the likelihood derived under Case (ii) would be relevant. However, it requires a complicated numerical maximization technique. The Expectation-Maximization (EM) algorithm might be applied in Case (ii). Further investigation on that case would be useful.

References

[1] Blischke, W. R. and Murthy, D. N. P. (2000). Reliability, Wiley, New York.

[2] Blischke, W. R., Karim, M. R. and Murthy, D. N. P. (2011). Warranty data collection and analysis, Springer.

[3] ISO 8402 (1994). Quality Vocabulary. International Standards Organization, Geneva.

[4] Meeker, Q. and Escobar, L. A. (1998): Statistical Methods for Reliability Data, John Wiley & Sons, Inc.

[5] Murthy, D. N. P., Xie, M., and Jiang, R., (2004). Weibull Models. Wiley, New York.