Statistical Process Control For Failure Crushing Time Data Using Competing Risks Model

F.A.M. Elfaki, J. Daud, M. Azram, Bin Daud, I, N.A. Ibrahim and M. Usman

Department of Science in Engineering, Kulliyyah of Engineering, IIUM Jalan Gombak, 50728, Kuala Lumpur, Malaysia.
Department of Mathematics, Faculty of Science and Environmental Studies University Putra Malaysia, 43400, Serdang, Selangor.
Faculty of Economics, Universitas Malahayati, Ji. Pramuka No. 27 Kemiling Bandar Lampung, Indonesia

Abstract: This paper describes a Statistical Process Control (SPC) for failure crushing time data using competing risks model. The model is based on the widely known proportional hazard regression model for a variety of censoring. A competing risks model identifies the set of possible failed components given the true cause of failure. EM algorithm method is used to estimate the parameter of the model. The results of this study show that, the competing risks model performs well for SPC using SAS software.

Key words: SPC; Competing Risks; Cox’s Model; Reliability; EM algorithm.

INTRODUCTION


The most important tool in SPC is the control chart and process capability index. There are two basic types of control charts: charts for variables data and charts for attributes data. Variable data control charts are useful when the parameter of interest can be conveniently measured numerically, for example, the measurement of the diameter of a cylindrical part. Whereas attribute data control chart are useful when the parameter of interest cannot be conveniently measured numerically, for example, the inspection of the finished surface of a cylindrical part. X̄ and R charts belong to the category of variable data charts, and p and c charts belong to the category of attribute data charts.

In this study, a traditional variable control charts commonly known as mean (X̄) and range (R) control charts were used to examine the state of control of the process. This will be first stage of SPC procedure. The variable control charts such as X̄ and R charts were chosen because it allows studying a process regardless of its specifications. The X̄ and R are also allow to employ both instantaneous variable (short-term process control charts) and long-term process control charts.
capability), and variability across the (long-term process capability). It is particularly helpful if the data for a process capability study are collected in two to three different time periods, such as different shift and different day’s. The variable control charts are important in the quality program of many industries, their ability to identify process improvement opportunities.

The theory of competing risks is applied in the analysis of reliability and survival data involving several different failure types or risks. In an industry, for instance, one might distinguish between a mechanical device failure attributable to a component that has failed and those due to unrelated causes. This constitutes the different risks under consideration. Typically, the data include the time of failure or censoring of each individual, as well as an indicator of the type of failures. To assess the effects of covariates on cause-specific hazards, one can perform a parametric Cox’s proportional hazards model, treating failure types which are of interest as censored observation (Aly, et al., 1994; Cheng et al., 1998).

In this paper, we use competing risks model that is, parametric Cox’s model with Weibull distribution based on EM algorithm, to examine the state of control of the process, that is, variable control charts, and estimation of capability indices. However, the parameters are obtained by using SAS software.

2. Statistical Process Control:

The goal of statistical process control (SPC) is to ensure that a process, with outputs that exhibit both systematic and random components, produces high quality items over time. To do so a process must be stable and consistent, with a steady systematic component and with a typical random component small enough not to seriously degrade item quality. (This is to say that the control limits should define an acceptable range of measurements to ensure adequate product quality). Because the random component causes the sequence of measurements to vary, process observers must try to distinguish between ordinary fluctuations (which require not intervention), and aberrant behavior (which must be corrected or at least understood). Control charts and process capability index based on competing risks model provide an objective way to do this.

Section 2.1 discusses control charts and section 2.3 discusses statistical process control concepts of capability.

2.1. The $\bar{x}$ Chart

The $\bar{x}$ control chart is a simple graphic procedure to study the state of the process. The control chart is a time-ordered sequences of observations of data of interest plotted between two horizontal lines called the control limits, the upper control limit (UCL), and lower control limit (LCL). Periodically, such as every hour, a sample of say $n$ items from the production process are picked and the sample average (denoted by $\bar{x}$) is plotted on the vertical axis. This process is continued and the consecutive points are joined by a straight line. If any point falls between the limits the process is said to be under statistical control. If a single observation lies beyond the control limits, the process is declared to be out of control and corrective action is taken to detect and eliminate the assignable cause. In general, the center line, the upper control limit, and the lower control limit are defined as:

$$UCL = \mu_i + L \frac{\sigma}{\sqrt{n}}$$

$$CL = \mu_i$$

$$LCL = \mu_i - L \frac{\sigma}{\sqrt{n}}$$

where $L$ is the distance of the control limits from the center line, expressed in standard deviation units. For example, the $\bar{x}$ chart, using $L=3$, would create a chart such the probability of type I error (i.e., the chart signaling “out of control” when the process is in control) occurring is 0.0027. In other words, an average of 27 false alarms is signaled out of 10,000 samples.
2.2 The R Control Chart:
Plotting values of the sample range \( R \) on a control chart may monitor process variability. The center line and control limits of the \( R \) chart are as follows:

\[
\begin{align*}
UCL &= D_4 \bar{R} \\
CL &= \bar{R} \\
LCL &= D_3 \bar{R}
\end{align*}
\]

The constants \( D_3 \) and \( D_4 \) are tabulated for various values of \( n \) Appendix of Montgomery (1989).

2.3. Process Capability:
Data built up over a period of using control charts in the way described may be used to know how the process ought to behave through the assessment of its capability, that is, its capacity to meet customer’s or other specifications. A review of historical control charts will establish the degree of statistical control of the process and may also be used as a basis for estimating mean \( \mu \) and standard deviation \( \sigma \). Following this, the ability of the process to meet specification is assessed through calculation of one or more capability indices. In practice, the values of \( \mu \) and \( \sigma \) usually are not known. The process must be stable in order to produce reliable estimates of \( \mu \) and \( \sigma \). Assuming that the process has reached a state of statistical control, the question often arises as to whether it can meet the tolerance.

The most commonly used measures of process capability are \( C_p \), \( CPU \) \( CPL \) and \( C_{pk} \). These indices have been utilized by a number of Japanese companies and in the U.S. automotive industry (Kane, 1986).

Several researches used process capability in their studies such as, Kane (1986) who describes a distribution of a function of \( C_p \). Specifically, the squared ratio of \( C_p \) to \( C_{pk} \) multiplied by (n-1) is Chi-squared distribution with (n-1) degrees of freedom. Kane also, develops sample size guidelines for \( C_p \) and critical value determination for the testing of \( C_p \). Clements (1989) proposes a way to evaluate process capability is not normal by using Pearson distribution curves. The data is transformed to a normal distribution and then \( C_p \) and \( C_{pk} \) are calculated from these transformations. While, Owen and Hua (1977) develop confidence limits in the tail areas of the normal distribution. Chou et al (1990) develop lower confidence limits on many process capability indices using paper published by Owen and Hua (1977) and they conclude that the distribution of \( C_{pk} \) follows a non-central t-distribution. In this paper we use the competing risks model to calculate a number of processes capability indices. A number of processes capability indices and their estimators will be present in the next sections.

3. Process Capability Indices and Their Estimators:
3.1 \( C_p \) Index:
The process capability index \( C_p \) is defined to be

\[
C_p = \frac{allowable\ process\ spread}{actual\ process\ spread} = \frac{USL - LSL}{6\sigma}
\]

where \( USL \), \( LSL \), and \( \sigma \) denote the upper specification limit, lower specification limit, and process standard deviation associated with the measurements, respectively. A process is said to be capable if the value of \( C_p \) associated with the process is at least 1.0 (Kane, 1986). Since the process standard deviation is rarely known, it is estimated from a sample of \( n \) measurements \( X_1, ..., X_n \) and an estimate \( \hat{C}_p \) of the process capability \( C_p \) is obtained by

\[
\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}}
\]

Typically, the sample standard deviation
The process capability index can also be considered as a measure of nonconforming product. An index value of one represents 2,700 parts per million (ppm) nonconforming, while 1.33 represents 63 ppm; 1.66 corresponds to 0.6 ppm; and 2 indicates fewer than 0.1 ppm. These values are correct if, and only if, the process measurement arises from a normal distribution centered on the midpoint of the specification limits. If this is not true, the process capability index will underestimate the percent nonconforming.

Next, several indices will be considered that take into consideration both the location of the process mean and the process variance. These indices reflect departures of the process mean from the target value and changes in the process variance.

Consider the two unilateral specification limit cases where only an upper or only a lower specification limit exists. The indices $CPU$ and $CPL$, which are used to monitor these cases, are defined below.

### 3.2 CPL and CPU Indices:

Similar to (7), the lower process capability index ($CPL$) is defined to be

$$CPL = \frac{\text{allowable lower spread}}{\text{actual lower spread}} = \frac{\mu - LSL}{3\sigma}$$

and the upper process capability index ($CPU$) is defined to be

$$CPU = \frac{\text{allowable upper spread}}{\text{actual upper spread}} = \frac{USL - \mu}{3\sigma}$$

Note that

$$C_p = (CPU + CPL)/2$$

The $CPU$ index was developed in Japan and is utilized by a number of Japanese companies (Kane, 1986).

### 3.3 Cpk Index:

We know that process variability is not the only parameter that influences a process’s ability to produce a conforming product. The location of process mean is another parameter that impacts process capability as suggested by Gunter (1989). Although we observed that one measure of process capability, the $C_p$ index, does not incorporate the process location, other indices do.

One index that accounts for this location, the $C_{pk}$ index, is used when the process mean is not at the target value, which is assumed to be halfway between the specification limits. The $C_{pk}$ index is given by

$$C_{pk} = \text{minimum (CPL, CPU)}.$$  \hspace{1cm} (13)

$C_{pk}$ describes a distance scaled by $3\sigma$, between the process mean and the closest specification limit. Assuming that $\mu$ is between the specification limits.

### 4. The Model:

The proportional hazards (PH) regression model is commonly used in the analysis of survival data and, recently, there has been an increasing interest in its application in reliability engineering. Following Cox’s (1972), we will focus on a particular model that is

$$h(t | z) = h_0(t) \exp(z\theta)$$

where $\theta = (\theta_1, \ldots, \theta_p)'$ is a vector of regression coefficients, $t$ is continuous random variable representing an individual’s lifetime, and $z = (z_1, \ldots, z_p)$ is vector of regressor variable associated with the individual.
The standard approach to inference for the two parametric regression models is the EM algorithm method. Here, if we observe a subject who failed at time \( t \), then, the contribution to the likelihood is \( f(t; \theta, z) \) the density function at \( t \). The contribution from a subject censored at \( t \) is \( R(t; \theta, z) \) the probability of reliability beyond \( t \). Thus, full likelihood based on the data \((t_i, \delta_i, z_i)\) \( i = 1,2,...,n \), is given by Kalbfleisch and Lawless (1988), and Lawless (1983), as follows:

\[
L(\theta) = \prod_{i=1}^{n} f(t_i; \theta, z_i)^{\delta_i} R(t_i; \theta, z_i)^{1-\delta_i} \tag{15}
\]

where \( \delta_i \)'s are event indicator variables: \( \delta_i = 1 \) if the \( ith \) subject fails; \( \delta_i = 0 \) if the \( ith \) subject is censored; \( \theta \) is a parameter that indexes the density function; and \( Z_i \) are the covariates for the \( ith \) subject.

Taking the natural logarithm of equation (15) simplifies the optimization. The log-likelihood function is given by Kalbfleisch and Lawless (1988), and Lawless (1983) as follows:

\[
l = \ln(L) = \sum_{i=1}^{n} \ln \left[ f(T_{F,i}) \right] + \sum_{j=1}^{S} \ln \left[ R(T_{S,j}) \right] \tag{16}
\]

where \( T_F \) is the exact time to failure and \( T_S \) is the censored time to failure. The model will be formulated in such a way that equation (16) will be a function of the stresses by expressing the probability density function (pdf) and reliability functions in terms of these stresses.

### 4.1 The PH Weibull Model:

The Weibull distribution is commonly used for analyzing lifetime data. Also, can be used as the underlying life distribution. In other words it is assumed that the baseline failure rate in equation (14) is parametric and is given by the Weibull distribution. In this case, the baseline failure rate is given by:

\[
h_0(t) = \eta \alpha^{-1}(t/\alpha)^{\eta-1} \exp \left[ -(t/\alpha)^\eta \right] \tag{17}
\]

where \( \alpha \) is the scale parameter depending on \( z \) and \( \eta \) is the shape parameter. In fact, \( \eta \) does not depend on implies proportional hazards for lifetimes and constant variance for log lifetimes of individuals. This assumption is reasonable in many situations, as discussed by Peto and Lee (1973), and Pike (1966).

The PH failure rate then becomes,

\[
h(t/z) = \frac{\eta}{\alpha} \left( \frac{t}{\alpha} \right)^{\eta-1} e^{\sum_{j=0}^{n} \theta_j z_j} \exp \left[ -t^{\eta} e^{\sum_{j=0}^{n} \theta_j z_j} \right] \tag{18}
\]

It is often more convenient to define an additional covariate \( Z_{\eta} = 1 \), in order to allow the Weibull scale parameter raised to the \( \beta \) (shape parameter) to be included in the vector of regression coefficients. The PH failure rate can then be written as:

\[
h(t/z) = \beta t^{\beta-1} e^{\sum_{j=0}^{n} \theta_j z_j} \exp \left[ -t^{\beta} e^{\sum_{j=0}^{n} \theta_j z_j} \right] \tag{19}
\]

The reliability function can be derived as,

\[
R(t,z) = e^{-\int_{z}^{t} \lambda(t,u) du} = e^{-\sum_{j=0}^{n} \theta_j z_j} \tag{20}
\]

where, \( \lambda(t, z) \) is failure rate of the model (14).
The pdf can be obtained by taking the partial derivative of the reliability function given by equation (20) with respect to time.

The reliability function and the Weibull pdf can then be substituted into equation (16). This yields the likelihood function for PHW model, as follows:

\[ l = \sum_{i=1}^{F} \ln \left( \beta T_{F,i}^{\beta-1} \exp \left( \sum_{k=0}^{m} \theta_{k} z_{i,k} \right) \exp \left( -T_{i}^{\beta} e^{\theta_{m}} \right) \right) - \sum_{k=1}^{S} T_{S,k}^{\beta} e^{\theta_{m}} \sum_{j=1}^{\theta_{j}} \]  

(21)

Solving the parameters that maximize equation (21) will yield the parameters for the PHW model, which are obtained by simultaneously solving the following partial derivatives:

\[ \frac{\partial l}{\partial \beta} = 0 \quad , \quad \frac{\partial l}{\partial \theta} = 0 \]

5. Failure Crushing Time Data (FCTD):

The proposed method is illustrated with some data taken from Mahdi (2003). The data consist of failure of woven roving wound laminated tubes. Initially the data were collected to assess the effects of mandrel rigidity on the load-carrying capacity and the energy absorption capability of woven roving wound laminated circular cylindrical composite shells design to withstand axial crushing.

A proportional hazards regression model with Weibull distribution, with three indicator variables, each representing a particular covariate, is fitted. Two data sets represent mandrel with internal diameters of 0 and 10 mm. 39 observation were taken for each data set. Failure loads at “diameter” 9.5 and 10.2 mm are observed giving 3 and 4 uncensored observations for wound on mandrel with internal diameter of 0 and 10mm respectively.

It is of interesting to note that the FCTD will be used for the statistical process control calculation based on competing risks models. Apart from that, the $\bar{X}$ chart and R chart, and process capability indices was also calculated for comparison between the three covariates (i.e. aluminum, plastic and wood) included in the FCTD.

**FCTD (d_i=0)**

It was observed that, the $\bar{X}$ chart and R chart from the first risk ($e=1$) seem to be in state of statistical control, for the aluminum and plastic covariates as in Figure 1, 2, 3 and 4. On one hand, the wood covariate from the same data has one point observed to be above upper control bound as clearly shown in Figure 5. On the other hand, Figure 6, which associated with $R$ chart, gives evidence of the presence of all the observation which are in state of statistical control. This results indicates that the tube wrapped on aluminum and plastic mandrels have along survival compared to the one wrapped on wooden mandrel. However, the results might be attributed to the second risk ($e=2$), which showed the tube wrapped on plastic mandrel have excellent crashworthiness performance with respect to the $\bar{X}$ chart and R chart as shown in Figure 9 and 10. on the other hand Figures 7 and 11 show the $\bar{X}$ chart for the aluminum and wood covariates. It can be seen that some observations found to be below the lower bounds of control limit. In contrast from $\bar{X}$ chart, $R$ chart gives an indication that the aluminum and wooded covariates are in the state of control as seen in Figure 8 and 12. Note that Figures 7, 8, 11, and 12 is not addressed here.

**FCTD (d_i=10)**
Fig. 1: The $\bar{X}$ control chart for aluminum mandrel with internal diameter of 0 mm (First Risk).

Fig. 2: The $\bar{X}$ control chart for aluminum mandrel with internal diameter of 0 mm (First Risk).

Fig. 3: The $\bar{X}$ control chart for plastic mandrel with internal diameter of 0 mm (First Risk).

Fig. 4: The $\bar{X}$ control chart for plastic mandrel with internal diameter of 0 mm (First Risk).
Fig. 5: The $\bar{X}$ control chart for wood mandrel with internal diameter of 0 mm (First Risk).

Fig. 6: The $\bar{X}$ control chart for wood mandrel with internal diameter of 0 mm (First Risk).

Fig. 9: The $\bar{X}$ control chart for plastic mandrel with internal diameter of 0 mm (Second Risk).

Fig. 10: The $R$ control chart for plastic mandrel with internal diameter of 0 mm (Second Risk).
For this data set (i.e. thick-walled mandrel). The $\bar{X}$ chart obtained from the two types of failure, for the aluminum, plastic and wood shifting below the control limit (LCL) which is indicate that to be out of control as seen clearly in Figure 13, 15, 17, 19, 21 and 23. On the other hand, the $R$ chart obtained from the two type of failure, for the aluminum, plastic and wood seem to be in state of control as seen in Figure 14, 16, 18, 20, 22 and 24 respectively. Note that Figures 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 and 24 is not addressed here.

Table 1 show the estimates of the $C_p$, $C_{PL}$, $CPU$ and $C_{pk}$, obtained by the competing risks model, that is, Cox’s model with Weibull distribution based on EM algorithm from two data sets of FCTD (0 mm and 10 mm). The value of $C_p$ and $C_{pk}$ for the first type of failure obtained from FCTD (d=0 and 10) are reasonably closed to one for covariates wood, plastic and aluminum. Moreover, due the two type of failure for same data set, plastic mandrel provide the estimated of the $C_p$ and $C_{pk}$ values are greater than one, which is indicate that the process is capable of producing items within desired limits.

6. Conclusion:

The parametric Cox’s proportional hazards regression model with Weibull distribution based on EM algorithm has been used successfully to investigate the causes of failure for statistical process control. The combination of the two data sets gives engineers the opportunity to perform analysis of more than one stress-type. Even though, from the analysis the first type of failure (stress) is significant, we have an insight of other causes that may can bute to the failure. The result obtained by proposed model for statistical process control technique in equations (1), (2), (3), (12), (13) based on equation (19) can be used effectively in the analysis of reliability data. Plastic mandrel shows that the process is under control but operating at an unacceptable level for covariates aluminum and wood. There is no evidence to indicate that the production of nonconforming units is operator controllable. It is quite clear to improve the energy absorption capability of woven roving composite tubes; they must be wrapped over a well design mandrel, such as the plastic one. Follow-up research should cover the application of these methods to simulation data.

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Covariate</th>
<th>$C_p$</th>
<th>$C_{PL}$</th>
<th>$CPU$</th>
<th>$C_{pk}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mm</td>
<td>Wood</td>
<td>0.99775</td>
<td>1.01182</td>
<td>0.98368</td>
<td>0.98368</td>
</tr>
<tr>
<td></td>
<td>Plastic</td>
<td>1.00222</td>
<td>1.00318</td>
<td>1.00126</td>
<td>1.00126</td>
</tr>
<tr>
<td></td>
<td>Aluminum</td>
<td>0.99834</td>
<td>0.99871</td>
<td>0.99797</td>
<td>0.99797</td>
</tr>
<tr>
<td>10 mm</td>
<td>Wood</td>
<td>0.99898</td>
<td>1.00036</td>
<td>0.99745</td>
<td>0.99745</td>
</tr>
<tr>
<td></td>
<td>Plastic</td>
<td>0.99423</td>
<td>1.00852</td>
<td>0.97995</td>
<td>0.97995</td>
</tr>
<tr>
<td></td>
<td>Aluminum</td>
<td>0.98064</td>
<td>0.98258</td>
<td>0.9787</td>
<td>0.9787</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Covariate</th>
<th>$C_p$</th>
<th>$C_{PL}$</th>
<th>$CPU$</th>
<th>$C_{pk}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mm</td>
<td>Wood</td>
<td>0.99865</td>
<td>1.01182</td>
<td>0.95386</td>
<td>0.95386</td>
</tr>
<tr>
<td></td>
<td>Plastic</td>
<td>1.00222</td>
<td>1.00318</td>
<td>1.00126</td>
<td>1.00126</td>
</tr>
<tr>
<td></td>
<td>Aluminum</td>
<td>0.98349</td>
<td>0.98719</td>
<td>0.97974</td>
<td>0.97974</td>
</tr>
<tr>
<td>10 mm</td>
<td>Wood</td>
<td>0.99802</td>
<td>1.00036</td>
<td>0.97459</td>
<td>0.97459</td>
</tr>
<tr>
<td></td>
<td>Plastic</td>
<td>0.99423</td>
<td>1.00852</td>
<td>0.97995</td>
<td>0.97995</td>
</tr>
<tr>
<td></td>
<td>Aluminum</td>
<td>0.97046</td>
<td>0.97582</td>
<td>0.9787</td>
<td>0.9787</td>
</tr>
</tbody>
</table>

REFERENCE


Dhillon, B.S., 1985. “Quality Control, Reliability, and Engineering Design.” Marcel Dekker, INC.


