CHAPTER EIGHT

CONCLUSION AND FURTHER WORKS

8.1. Summary of the Study

This study is aimed at developing influential observations detection for Cox PHM. We first studied the residual that is being used to identify outliers in Cox PHM. Three types of residual are used; deviance, normal deviate and log-odds residuals. A simulation study had been carried out to investigate the sampling distribution of the residuals in the presence of censored observations. We showed that as the percentage of censoring increases, the residuals failed to follow their respective sampling distributions. Hence, we studied the sampling behavior of the maximum or minimum values of the residuals which can then be used as the cut points for outlier detection purposes. Using three real data sets as illustration, we showed that normal deviate and log-odds residuals gave better identification of outliers compared to deviance residual.

Then, these three residual were used in nine different new methods to identify influential observation. This is called the Cox PHM FS method, denoted by FS1 till FS9. The performance of these methods also had been illustrated using three survival data sets. In general, methods based on the normal deviate and log-odds residuals perform better in identifying influential patients than that based on the deviance residuals. The results also suggest that the new methods improved the detection of influential observations if compared to the delete-case method.

8.2. Significance of the Study

This study focused on the improvement of outlier and influential observation detection procedures in Cox PHM. It has contributed to the survival analysis in the following ways:

1. The two new types of residuals which were r_{N_i} and r_{L_i} gave better results in detecting outlier for Cox PHM.
2. The nine proposed Cox PHM FS methods showed a good improvement in detecting influential observations for the Cox PHM compared to the delete-case method.

8.3. Further Research

This research can be extended in many ways. As we know, the presence of the influential observations in survival analysis may give undue effect to the parameter estimate, where in this study the partial likelihood estimate was used. Therefore, the outlier and influential observation detection method can be extended in its application on other approaches to parameter estimate such as bootstrap, EM algorithm or maximum spacing techniques.

On the other hand, this method can also be extended to other types of survival model, such as Weibull or Stratified Cox PHM based on detecting outliers and influential observations. In this study we observed the single outliers; however the Cox PHM FS method can further detect “masked” outliers or influential observations in the data set. This method also can be further applied to identify observations that do not belong to the data set.