|  |
| --- |
| School |
| DESaster project |
| Name |

This report will cover the estimation of probability distribution using the data from Seattle government website with the permit for constructions. The estimated probability will be used in the DESaster Application Template file to redefine the distribution for time parameters, and outputs will be compared in this report as well.

**Introduction**

DESaster is the discrete event simulation for the disaster recovery process. It provides simulations in different stages and period after the disaster happened. The policies for recovery process may various depends on situation, but in general there will be time frame between each stage for inspection, application, getting permit and so on. The days between each stage are essential for estimate the complete cycle for the recovery process. This is essential to individuals that need to repair their houses after disaster to know the estimate time spend on each stage to be prepared, also it will help on the insurance agencies to estimate the recovery funding. In the same time, the estimate duration could help on planning the repair process.

In the template provided on GitHub, all the durations are assumed to be normally distributed, but it seems like that there are no validations in the simulation model to check the fitness of the distribution. Theoretically, the more fit between assumption and the real distribution, the more accurate and reliable simulation results will be generated. This project aims to help on verify if some time parameters follows the normal distribution. If there is any probability distribution that fits better than normal distribution, the assumptions in the template will be replaced by the better distribution, and then the output, especially the figures shows time frame after disaster in the outputs to compare the differences before and after the distribution assumption change.

The main supporting material for this project is the documentation from GitHub of DESaster, some supplemental materials are found online for analysis on the fitting problems. All the reference documents will be cited in text and at the end of report.

**Methods**

**Variable description**

In the empirical data file, only the date parameters will be used in the estimation of probability distribution. The date parameters include Application Date, Issue Date, Final Date and Expiration Date. Application Date indicates for the date that application was accepted as a complete submittal. Issue Date indicates for the date that the application was issued as a valid permit. Final Date indicates for the date the permit had all its inspections completed. Expiration Date indicates for the date that the application is due to expire, which is the date that the work is supposed to be completed. Parallel to the parameters in disaster application template, the difference between Issue Date and Application Date will be the start\_delay\_dist; the difference between Final Date and Issue Date will be the inspection\_dist; the difference between Expiration Date and Final Date will be the repair\_dist.

**Model creating**

First, the data points will be organized that only the date remains.

Second, since there are some missing values inside the data file, but the missing values indicates the process are still ongoing or not provided, so the missing values will be removed.

Third, the data points will be plotted and fitted in figure. There may not be an exact fit, but instead, the distribution that with the lowest p-value will be identified for the best fit. The fit distributions will include normal, beta, uniform and Weibull.

Finally, the best fitted probability distribution will replace the normal assumptions which state in the disaster application template.

The results will be compared for the original normal assumption and the replaced distribution. Any significant changes detected in this process could possibly indicate that the normal assumption may not be the best fit for that parameter.

**Results**

**Estimated probability distribution**

**Tables and Figures**

**Discussion**

**Implementation**

As mentioned in introduction, the exploration of duration between every two stages will be helpful for all people involved in the disaster recovery process.

**Recommendations**

First, clearly the empirical data file is only limited to one city, therefore it is not a general study. Second, the durations for construction, inspection and other stages in disaster recovery may have many external factors. In this model, only few time parameters are examined and performed. More time parameters with more specific reference could be recorded to help on studying the best fitted distribution in future studies. Also, to fit the simulation model better, the actual recovery process and duration could be monitored and estimated to give the more accurate prediction on the outputs. Besides, we may study on the correlation among each period to estimate the complete recovery cycle under different policies in a more realistic way.

**References**

DEsaster, Scott Miles, updated April 2018

<<https://github.com/milessb/DESaster/blob/master/images/AboutDESaster10312017.pdf>>