**AA502 – Survival Analysis**

**Hurricane HW2**

Consider the following “hypothetical” situation. Several hurricanes struck the gulf area and resulted in severe casualty and property damage. One of the major defenses is to maintain and coordinate the pump operations during a critical 48 hour period or over 4 high tides. The Steering Committee of the Center for Risk Management is conducting a survival analysis for the pump stations in the gulf coast area. Data was collected from 771 pump stations. There are five potential failure conditions during a major hurricane hit recorded in the variable **reason**:

1. No failure
2. Flood – overflow or accumulation of an expanse of water that submerges the pump station.
3. Motor – mechanical failure.
4. Surge – onshore gush of water usually associated with levee or structural failure.
5. Jammed – accumulation of trash or landslide materials.

The factors that potentially influence the survivability of the pump stations are the following (not all pumps have each characteristic, but some characteristics are available through upgrade or maintenance):

* **Backup pump** – a redundant system used to protect the station from flooding when the main pump is not operating. (Upgrade available)
* **Bridge Crane** – allow vertical access to equipment and protecting materials. (Upgrade available)
* **Servo** – servomechanism is used to provide control of a desired operation through the Supervisory control and data acquisition (SCADA: Supervisory Control And Data Acquisition) systems. (Upgrade available)
* **Trash-rack cleaner** – protecting hydraulic structures against the inlet of debris, of vegetation, urban or industrial trash. (Upgrade available)
* **Elevation** – elevation of the pump station that can be altered by 1 foot by maintenance. (Maintenance available)
* **Slope** – surrounding ravine slope of the pump station.
* **Age** – difference between the installation and the current date.
* **H1 ~ H48** – pumping status during a 48 hour emergency reported by pump stations – accuracy of pump status not guaranteed to be error free.

As the Steering Committee of the Center for Risk Management, provide a follow-up report to your last report and set of recommendations summarizing the findings from the survival analysis performed on the data. The new follow-up report should include the following information:

* This analysis will only be focused on one type of failure (Flood), but could be repeated for all types.
* **Use an alpha = 0.03 for this entire analysis**.
* Use PROC LIFETEST to plot the following graphs. Instead of using the **survive** variable as your censor variable, use the **reason** variable with the values of 0, 2, 3, 4 all as censored (put these values in your parentheses separated by commas). This method is called competing risks, which we will talk about in the final class. Plot three graphs:
  + The survival curve
  + The log of the survival curve by time.
  + The log of the log of the survival curve by the log of time.
  + By these plots, can you tell anything about the possible distribution of your data?
* Using PROC LIFEREG, run AFT models with the same censoring as mentioned above with all of the following variables in your model: **backup**, **bridecrane**, **servo**, **trashrack**, **elevation**, **slope**, and **age**. Do this for four distributions: Exponential, Weibull, log-Normal, and Gamma.
* Using Likelihood Ratio Tests, discuss anything about the possible distribution of your data.
* Once you choose a distribution for your model, reduce your AFT model down to only having significant variables (do not worry about interactions in this assignment). Output these parameter estimates for later use.
* Create a new data set that **only** contains pumps that did **not** survive the hurricane due to flood. Use a similar method to the one we did in class with the prisoner recidivism data to do the following:
  + You can only perform one upgrade on each pump and your model should have two variables that could be upgraded.
  + In a new data set, isolate out the pumps that didn’t have either of these upgrades.
  + Calculate the average, median, minimum, and maximum predicted time difference that the pumps would have survived if they did have one upgrade.
  + Repeat this process for the other possible upgrade.
  + If you could only tell the Army Corp of Engineers one upgrade to make across all pumps, what upgrade would it be? Assume the upgrades both cost the same amount.