STA 222: Survival Analysis Final

Due Dec. 12th, Wed, 5pm.
Please hand in your hard copy of report in TA's mailbox. (Main office, 4th floor)
Use any software you feel comfortable with. Some **R** tips are given.
Make sure you answer the fill-in question. :)

1. Generate the data set
   - Generate $\{T_i\}_{i=1}^n$ following the hazard rates
   $$\lambda_i(t) = e^{\beta Z_i}\lambda_0(t)$$
   - Specify the baseline hazard you use and how you generate the covariates $\{Z_i\}_{i=1}^n$. Report the true parameter β.
   - Generate the censoring time $\{C_i\}_{i=1}^n$. Report the distribution you use.
   - Report the observed data $\{X_i, \delta_i, Z_i\}_{i=1}^n$.
2. Cox proportional hazard model
   - Calculate and report the MLE $\hat{\beta}_{PH}$ following Cox's conditional likelihood. Calculate and plot the GMLE of $\Lambda_0(t)$. Compare and discuss the estimations with your true parameter and baseline hazard.
   - Construct a goodness-of-fit test on the model. Verify and report your null hypothesis. Does the test-statistics you get indicate a good model fitting?
3. Accelerated Failure Time model
   - Fit the AFT model
   $$\log(T_i) = \beta Z_i + \epsilon_i$$
   and report the Buckley-James estimation $\hat{\beta}_{BJ}$ Is the value far from the estimator $\hat{\beta}_{PH}$? Discuss it.
   - Discuss the goodness-of-fit on the model.

A couple of tips in programming in **R**:

- The proportional hazard estimator is available in **R** package "survival", which is a standard package people used. Function *coxph* provides you the information related to the PH estimator. Make sure you check the help file before using it.
- Buckley James estimator are available in two **R** packages: "rms"(function bj) and "emplik" (function $BJnoint$). Be careful with the model settings when using them. In the first package, it even provide you the variance estimation. You are welcomed to explore them.