

Exam 2

Steven Chiou

Instructions

- Due date: **Tuesday, December 12.**
- This is a open resource exam, but you are not allowed to ask post exam questions online.
- You are not allowed to collaborate with classmates and/or people outside of class.
- Please circle or highlight your final answer.
- The total possible point is xx.

Violation of this agreement will result in an **F** on this exam and it will be averaged in as a 0%.

-
1. Under the exponential model, we assume T follows an exponential distribution with parameter λ and derived the maximum likelihood for λ to be

$$\hat{\lambda} = \frac{\sum_{i=1}^n \Delta_i}{\sum_{i=1}^n t_i},$$

where Δ_i 's is the censoring indicator and t_i 's are the observed survival times. Now let's suppose each copy of $\{\Delta_i, t_i\}$ is associated with a weight w_i (this weight could represent sampling weights, propensity score, or counts, etc.).

- a. Modify the likelihood presented in note 3 and derived the weighted maximum likelihood estimator for λ . We will denote this $\hat{\lambda}_w$.
 - b. Use `los` (length of stay) as the weight in *WHAS100*. Apply `survreg` to *WHAS100* to compute $\hat{\lambda}_w$.
 - c. Use the derivation in 1a and *WHAS100* to compute $\hat{\lambda}_w$.
2. In exam 1, we have investigated different methods to compare two survival curves. Another way to compare two survival curves is to fit a Cox model using the group indicator as the covariate. Use the complete *WHAS100* dataset and gender as the group indicator for the following questions.
 - a. Fit a Cox model and print the summary using gender as the covariate.
 - b. Interpret the estimated regression parameter ($\hat{\beta}$) in terms of hazard ratio.
 - c. Use part 2a to test for the null hypothesis that there is no significant difference between the two survival curves.
 3.
 - a. Fit a Cox model with `age` as the (only) covariate. Plot the estimated cumulative hazard function for patients with `age = 50`, `age = 60` and `age = 70`.
 - b. Fit another Cox model with `age` and `age^2` as covariates. Plot the estimated cumulative hazard function for patients with `age = 50`, `age = 60` and `age = 70`.
 - c. Interpret the results in 3a and 3b.