Homework 2 (Due, Friday 04/21)

1. Suppose the observed data are

\[ \{(U, \delta) = (1, 1), (1, 0), (3, 0), (4, 1), (5, 1), (5, 0), (9, 0), (12, 0)\} . \]

Calculate the KM estimator. If the last observation is (12, 1) instead of (12, 0), what is the corresponding KM estimator?

2. The variance of KM estimator \( \hat{S}(t) \) can be estimated by the Greenwood formulae:

\[
\hat{S}^2(t) \sum_{i=1}^{j} \frac{d_i}{Y(\tau_i)(Y(\tau_i) - d_i)}
\]

for \( t \in [v_j, v_{j+1}) \), where \( 0 < v_1 < v_2 < \cdots < v_p \) are all the observed failure times.

(a) Give a heuristic construction of an estimator for the covariance

\[
\text{cov}(\hat{S}(t), \hat{S}(s)),
\]

where \( t \in [v_j, v_{j+1}) \) and \( s \in [v_k, v_{k+1}) \) with \( 1 \leq j < k \leq p - 1 \).

(b) Construct a 95% confidence region for \((S(t), S(s))\). (Note that the construction is not unique and you only need to present one; Hint: consider the joint distribution of \{\(\hat{S}(t) - S(t), \hat{S}(s) - S(s)\}\).)