PubH 7450 Homework 2 (Spring 2023) Due date: March 30, 2023

Steffen Ventz

2023-03-22

Note: Please use an Rmarkdown document (Rmd and pdf) which includes your R code; show major steps in your derivations/calculations.

1) [10pts] The ''tongue" dataset (available in the KMsurv package) reportes data from a study that evaluates the effect of ploidy on survival for patients with tumors of the tongue (see Sec 1.11 in the KM-book).

(i) Test the hypothesis that the survival rates of patients with cancer of the tongue are the same for patients with aneuploid and diploid tumors using the log-rank test.

(ii) Assume that the primary interest is in detecting differences in survival rates between the two types of cancers which occur soon after the diagnosis of the cancer. Repeat part (i) using a more appropriate test statistic. Note: you may use the R package "FHtest'' for this test.

(iii) For the outlined testing problem on (ii) use instead a two-sample test based on the 56-weeks RMST.

2) [5pts] For this excercise use the data on laryngeal cancers ("larynx'' in the KMsurv package, see also Example 7.6 in the KM-book).

(i) Test, by the log-rank statistic, the null hypothesis of no difference in death rates among the four stages of cancer against the global alternative that at least one of the death rates differs from the others. Compare your results to those found in Example 7.6 in the book.

3) [10pts] The data on laryngeal cancer patients was collected over the period 1970–1978. It is possible that the therapy used to treat laryngeal cancer may have changed over this nine year period.

(i) To adjust for this possible confounding fact, test the hypothesis of no difference in survival between patients with different stages of disease against a global alternative using a test which stratifies on the cancer being diagnosed prior to 1975 or not.

(ii) Also perform a separate test of the hypothesis of interest in each stratum.

4) [10pts] The "kidtran'' data (in the KMsurv package, see also see section 1.7) report on the death times of 863 kidney transplant patients. Here, patients can be classified by race and sex into one of four groups.

(i) Test the hypothesis that there is no difference in survival between the four groups (all combinations of race and sex).

(ii) Provide individual tests, for each sex, of the hypothesis of no racial differences in survival rates.

(iii) Also, adjusting by stratification for the sex of the patient, test the hypothesis that blacks have a higher mortality rate than whites.

5) [10pts] Use the Bone marrow transplant data ("bmt'' available in the KMsurv package, see also Section 1.3 in the KM-book) to

- (i) compare the three survival functions for ALL, AML low-risk, and AML high-risk and test the hypothesis that the survival rates of these 3 groups are the same.

- (ii) Repeat the analysis in (i) using a stratified test and stratify for recruitment site (i.e. Hospital)

- (iii) Perform perform pairwise tests to compare the three survival functions for ALL, AML low-risk, and AML high-risk at one year. Use a Bonferroni correction (see Sec 7.8, p. 237) for multiple tests.