# BIOS 622 Homework 5

Instructor: Sijin Wen

March 1, 2025

Please answer each question precisely and completely. Use clear, complete sentences when giving your answers and derivations. It is permissible to discuss each other, but all of your work must be your own.

1. This homework is the nonparametric analysis. That is, we will not make the assumption of exponential distributions. A modified example from the course notes (Page 60): the data of time to death from severe viral disease $(T_1, ..., T_n)$ are given as follows:

   - Steroid: 1(2), 1+(2), 4+, 5, 7, 8, 10, 10+, 12+, 16+(3)
   - Control: 1, 2, 3(2), 3+, 5+(2), 16+(8)

   For each $t_i$:

   $D_i$ = Number of events (deaths) at time $t_i$

   $S_i$ = Number that are known to have survived beyond $t_i$ (by convention: includes those censored at $t_i$)

   $N_i$ = Number of "at risk" of being observed to die at time $t_i = S_i + D_i$

   $S(t_i)$ = survival function at $t_i$

   $h_i$ = hazard function at time $t_i$

   $H_i$ = corresponding cumulative hazard

   (a) Find Kaplan-Meier estimator of survival function for each group (with standard error) by completing the following table:

| $i$ | $t_i$ | $N_i$ | $D_i$ | $S_i$ | $\hat{S(t_i)}$ | s.e.$(\hat{S}(t_i))$ | $\hat{H}_i$ | $\hat{h_i}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| . | . | . | . | . | . | . | . | . |
| k | | | | | | | | |

(b) Plot survival curve (using eithor R or SAS) for each group based on Kaplan-Meier method. What is median survival in each group? Please also include the output from R and SAS which should give you the same result as in (a).

(c) Plot survival curve (eithor R or SAS) for each group based on Life-table. Is there any difference between approaches (b) and (c)?

(d) What is the p-value from log-rank test between two groups? Please state the null and alternative hypotheses.

(e) (Bonus) Can you estimate the mean survival (a non-parametric method) for each group based on the table in (a)?