Fall, 2005 October 18, 2005
Instructor: Pingfu Fu

Midterm Exam
EPBI 435: Survival Data Analysis

This exam is closed book and closed notes. There are 3 questions plus a bonus question and the duration of the exam is 90 minutes. Please show all work, but be brief and precise in your answers. The relationship between survival function $S(t)$ and cumulative hazard function $H(t)$ is $S(t) = exp(-H(t))$ in case you need it.

1. To compare two treatments for solid tumors, two groups of mice were inoculated with a highly malignant neoplasm, and after randomization into the treatments, survival (days) was observed as follows:

```
        Group 1, drug A: 16, 16+, 18+, 19+, 20, 28, 32+
        Group 2, drug B: 10, 14, 15, 18, 18, 20+, 21
where "+" indicates censoring (40 points).
```

   (a) Compute the Kaplan-Meier product-limit estimate of each group and plot the two survival curves in the same figure [hint: K-M estimator is $\hat{S}(t) = \prod_{j=1}^{k}(1 - d_j/n_j)$ for $t_{(k)} \leq t < t_{(k+1)}$].

   (b) Find the median survival time for each group.

   (c) Compute the cumulative hazard at 20 days for each group. Which drug is more effective? Why?

   (d) What does d_j/n_j in the formula of K-M estimator estimate? Give your answer as a conditional probability. Comment briefly how information from the censored observations contributes to the K-M estimator.

2. Cox proportional hazards models (30 points). Consider two groups of survival data with hazards $\lambda_1(t)$, $\lambda_2(t)$ and survivor functions $S_1(t)$, $S_2(t)$, respectively.

   (a) One of the assumptions for Cox model is proportional hazard, what is really meant by "proportional hazard"?

(b) Assuming the two hazard functions are not the same, examine the connection between the crossings of the two hazard functions and the crossing of the two survivor functions.

3. Averbook (2002) reports data on 1018 patients with resectable malignant melanoma during the period 1970-2000 at MetroHealth Medical Center in Cleveland. Reference: Averbook BJ et al (2002). Surgery, Vol 132, pp 589-604.

   The variables in the final Cox model are as follows:

```
THICK: tumor thickness (millimeters); ULCER: ulceration (yes = 1, no = 0)
AGE:   age at surgery (year)
TYPE: tumor type (0 for superficial spreading, 1 for nodular)
st: Time to death; censor: censoring indicator (0=alive, 1=dead)
```

   Based on the results generated from SAS procedure PROC PHREG, answer following questions (40 points)

   (a) Find out the hazard ratio of 10 millimeters increase of tumor thickness.

   (b) For the following two cohorts of patients with the same age and the same tumor thickness:
   (I) patients with ulceration, spuperficial spreading tumor type.
   (II) patients without ulceration, nodular tumor type.

      i. What is the hazard ratio between the two cohorts ?
      ii. Use delta method to calculate 95% confidence interval of above hazard ratio.

4. Optional (bonus): Let $T_1, T_2, ..., T_k$ be k independent continuous non-negative random variables with hazard function $\lambda_1(t), \lambda_2(t), ..., \lambda_k(t)$, respectively. Show that the hazard function for the random variable $T = min(T_1, T_2, ..., T_k)$ is $\lambda_T(t) = \lambda_1(t) + \lambda_2(t) + ... + \lambda_k(t)$. (10 points).