**BIOSTAT/STAT 576**
**Statistical Methods for Survival Data**

**Problem Set 1**
**Due: April 7, 2009**

1. Consider the life table used in Lecture 1:

| Years | Alive at beginning | Deaths | Censored |
|---|---|---|---|
| 0-1 | 146 | 27 | 3 |
| 1-2 | 116 | 18 | 10 |
| 2-3 | 88 | 21 | 10 |
| 3-4 | 57 | 9 | 3 |
| 4-5 | 45 | 1 | 3 |
| 5-6 | 41 | 2 | 11 |
| 6-7 | 28 | 3 | 5 |
| 7-8 | 20 | 1 | 8 |
| 8-9 | 11 | 2 | 1 |
| 9-10 | 8 | 2 | 6 |

   (a) Complete this ten-year life table by calculating survival rate for each time interval, assuming that censoring occurred at

   - the right end of each time interval,
   - immediately prior to the left end of each time interval, and
   - immediately prior to the left end of each time interval for half of the censored participants and at the right end of each time interval for the other half of the censored participants,

   respectively.

   (b) Plot the life table estimates of survival functions and their point-wise confidence intervals, respectively. Comment on your plots.

2. Show the inversion formula of survival function $S(t) = \Pr\{T > t\}$.

   (a) Let $m(t) = E(T - t \mid T > t)$ for a continuous nonnegative random variable $T$ with finite mean. Then

$$S(t) = \frac{m(0)}{m(t)} \exp\left[-\int_0^t \frac{1}{m(u)} du\right].$$

(b) (Kalbfleisch & Prentice, p. 29, Exercise 1.7) Let $m_i = E(T - i \mid T > i)$ for an integer-valued random variable $T > 0$ with finite mean $m_0$. Then

$$S(t) = \prod_{i=1}^{t} \frac{m_{i-1} - 1}{m_i}.$$

3. Suppose that the collected data are $\{(X_i = \min(T_i, C_i), \Delta_i), i = 1, 2, \ldots, n\}$, with $\lambda_{T_i}(t) = \lambda(t)$. Let $N_i(t) = I(X_i \le t, \Delta_i = 1)$ and $\Delta_i = I(X_i \ge t)$. Assume $T_i$ and $C_i$ are independent.

   (a) Sketch a proof to show that $\hat{\lambda}(t)dt = \sum_i dN_i(t) / \sum_i Y_i(t)$ is a consistent estimator of $\lambda(t)dt$, and state what assumptions you may need for your proof.

   (b) Show that $m(t)\lambda(t)dt = dm(t) + dt$. Then find a proper estimator for $m(t)$ using $\hat{\lambda}(t)dt$.

4. (Kalbfleisch & Prentice, p. 50, Exercise 2.2)

   (a) Show that the exponential distribution is the only continuous distribution for which $m(t)$ is constant for all $t > 0$.

   (b) Show that if $\lambda(t) \to \lambda$ as $t \to \infty$, then $m(t) \to 1/\lambda$.

   (c) Examine the form of $m(t)$ for log-normal and gamma distributions.

5. Suppose that the collected data are $\{(X_i = \min(T_i, C_i), \Delta_i), i = 1, 2, \ldots, n\}$. Consider a Weibull model for $T_i$ with $f(t) = \lambda\gamma t^{\gamma-1} \exp(-\lambda t^{\gamma})$.

   (a) Derive the score equations and the Fisher information for $(\lambda, \gamma)$

   (b) Denote the MLEs of $\lambda$ and $\gamma$ by $\hat{\lambda}$ and $\hat{\gamma}$, respectively. Derive a standard error estimator of the median survival time $\hat{t}_{0.5} = (\log 2/\hat{\lambda})^{1/\hat{\gamma}}$.

   (c) Propose a simulation plan to assess the performance of $\hat{t}_{0.5}$ you derived.