Midterm
(Time allowed: 1 hour 40 minutes)

Name: Solutions

INSTRUCTIONS

1. Write your Full name clearly in the space above.

2. This paper comprises Six (6) printed pages. Answer ALL the questions in the spaces provided.

3. This is a close book exam but students are allowed to bring TWO 2-sided help sheets of letter size.
Section A (5 marks)

1. If, for a given individual, the event of interest has already occurred when observation begins, then this individual’s survival time is truncated.
   ☒ TRUE ☐ FALSE

2. If, for a given individual, the event of interest was not observed when the study ended, then this individual’s survival time is right-censored.
   ☒ TRUE ☐ FALSE

3. Individuals who experience the event of interest at time $y$ are at risk at time $y$.
   ☒ TRUE ☐ FALSE

4. The hazard function ranges between 0 and $\infty$.
   ☒ TRUE ☐ FALSE

5. The hazard function at time $y$ gives the conditional probability for an event of interest to occur at time $y$, given survival up to time $y$.
   ☐ TRUE ☒ FALSE

6. The only restriction on the survival function is that it be nonincreasing.
   ☐ TRUE ☒ FALSE

7. Nonparametric methods do not require specific assumptions to be made about the underlying distribution of the survival times. Therefore, the Kaplan-Meier estimator of a survival function does not require the assumption of noninformative censoring.
   ☒ TRUE ☐ FALSE

8. The Kaplan-Meier estimate of a survival function is constant between adjacent (uncensored) survival times.
   ☒ TRUE ☐ FALSE

9. If the largest observed time corresponds to an uncensored survival time, say $y^*$, then the Kaplan-Meier estimate is zero beyond $y^*$.
   ☒ TRUE ☐ FALSE

10. If the largest observed time corresponds to a censored survival time, say $y^*$, then the Kaplan-Meier estimate is zero beyond $y^*$.
    ☐ TRUE ☒ FALSE
Section B (20 marks)

1. Suppose that a survival random variable $Y$ has probability density function (an average of two exponential distributions)

$$ f(y) = 0.5\lambda e^{-\lambda y} + 0.5\alpha e^{-\alpha y}, \quad y > 0. $$

(a) Find the survival function for $Y$. [2 marks]

(b) Find the hazard function for $Y$, $h(y)$. Under what condition, will $h(y)$ be a constant function? [3 marks]

\(\textbf{c\text{.a) } S(y) = P[ Y > y ]}\)

\[ = \int_{y}^{\infty} .5\left(\lambda e^{-\lambda t} + \alpha e^{-\alpha t}\right)dt \]

\[ = \left[ .5\left( e^{-\lambda t} + e^{-\alpha t}\right) \right]_{y}^{\infty} = .5\left( e^{-\lambda y} + e^{-\alpha y}\right), \quad y > 0.\]

\(\textbf{c\text{.b) } h(y) = \frac{f(y)}{S(y)} = \frac{.5\left(\lambda e^{-\lambda y} + \alpha e^{-\alpha y}\right)}{.5\left( e^{-\lambda y} + e^{-\alpha y}\right)} }\)

\[ = \frac{\lambda e^{-\lambda y} + \alpha e^{-\alpha y}}{e^{-\lambda y} + e^{-\alpha y}}, \quad y > 0.\]

\[ \text{when } \lambda = \alpha, \quad h(y) = \lambda = \alpha, \text{ a constant function.} \]
2. Censoring and Likelihood Construction

(a) [4 marks] The following data resulted from a clinical trial to compare treatments for the inhibition of relapse after an initial ulcer has been diagnosed, treated and healed. Regular visits to a clinic were arranged for the patients, and endoscopies were performed 6 months and 12 months after randomization. A positive endoscopy result (result = 2) indicates that an ulcer has recurred in the time since the last negative result. A negative result (result = 1) indicates that the patient is still in remission. Additionally, endoscopies were performed at in-between times (that is, between scheduled visits) for patients suffering from symptoms of recurrence. Such patients with positive test results provide 'exact' observed recurrence times.

Construct the censoring times for the censored patients. Clearly indicate the type of censoring in each case. For uncensored cases, say “uncensored” and leave “Censoring time” blank.

<table>
<thead>
<tr>
<th>Time of Last Visit (months)</th>
<th>Result</th>
<th>Censoring time</th>
<th>Type of Censoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>2</td>
<td></td>
<td>uncensored</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>12</td>
<td>right censored (Type I)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>right censored (random)</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
<td>left censored</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td></td>
<td>uncensored</td>
</tr>
</tbody>
</table>

(b) [2 marks] Write down the likelihood for this portion of the study.

\[
L \propto f(7) f(10) S(12) S(6) (1 - S(6))
\]
3. The following table determines the Kaplan-Meier estimate and the corresponding estimated standard errors of a right-censored survival data.

<table>
<thead>
<tr>
<th>j</th>
<th>t_{(j)}</th>
<th>n_{j}</th>
<th>d_{j}</th>
<th>\hat{S}(t_{(j)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>8</td>
<td>1</td>
<td>(1) .875</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>7</td>
<td>1</td>
<td>(2) .950</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>5</td>
<td>2</td>
<td>0.45</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>2</td>
<td>1</td>
<td>(3) .225</td>
</tr>
</tbody>
</table>

(a) Fill in the missing entries (1),(2) and (3). Show your work. [3 marks]

\[ \hat{S}(t_{(k)}) = \prod_{j=1}^{k} \left( 1 - \frac{d_j}{n_j} \right) \]

(1): \( 1 - \frac{d_1}{n_1} \) = \( 1 - \frac{1}{8} \) = .875

(2): \( (1) \times (1 - \frac{d_2}{n_2}) \) = .875 \times (1 - \frac{1}{7}) = .750

(3): \( .45 \times (1 - \frac{d_3}{n_3}) \) = .45 \times (1 - \frac{1}{2}) = .225

(b) Sketch the Kaplan-Meier survival curve. [2 marks]
(c) From the table, can we tell how many subjects were censored? Identify the time
intervals \([t_{(j)}, t_{(j+1)})\) of these censored observations. \([2 \text{ marks}]
\]

\[
\omega = n - d = 8 - 5 = 3
\]

\[
(23, 31), (31, 48), (48, \infty) \\
(\infty, t_{(3)}), (t_{(3)}, t_{(4)}), (t_{(4)}, \infty)
\]

(d) Denote the censored observations as \(c_1, c_2, \) and so on. Write the likelihood function
of the eight observations. \([2 \text{ marks}]
\]

The 8 observations:

18, 23, c1, 31, 31, c2, 48, c3

\[
L \propto f(18)f(23)SC(c_1)f(31)^2SC(c_2)f(48)SC(c_3)
\]