MATH 2625: Biostatistical Methods
Homework 5, due Thursday, April 10

Please submit a PDF or .doc version of your homework to Canvas by 3:30pm on the due date. Please type *all* responses. You are encouraged to use `R` for all calculations.

Theory

1. Find the estimated parametric survivor, hazard, and cumulative hazard functions for each of the following distributions. Because these distributions require root finding algorithms to find the MLE for some of their parameters, we will fix those parameters at the values specified below. Please show all of your work.

   (a) $T_i \sim Pa(\mu, \alpha)$ for $i = 1, \ldots, n$, $\mu = T_{(1)}$ (i.e. the minimum survival time), and α unknown.

   (b) $T_i \sim W(\lambda, \gamma)$ for $i = 1, \ldots, n$, $\gamma = 2$, and λ unknown.

Proof. The pdf of $T_i \sim Pa(\mu, \alpha)$ is defined as

$$f(t) = \frac{\alpha \mu^\alpha}{t^{\alpha+1}}.$$

We can integrate $f(t)$ to find the cdf, $F(t)$.

$$\begin{aligned}
F(t) = \int f(t)dt &= \int \frac{\alpha \mu^\alpha}{t^{\alpha+1}} dt \\
&= \alpha \mu^\alpha \left(\frac{1}{\alpha t^{-\alpha}} \right) \\
&= \left(\frac{\mu}{t}^\alpha \right).
\end{aligned}$$

Then our survival function $S(t)$ is defined as

$$S(t) = 1 - \left(\frac{\mu}{t}^\alpha \right).$$

The hazard function $h(t)$ of the Pareto distribution is then

$$h(t) = \frac{\frac{\alpha \mu^\alpha}{t^{\alpha+1}}}{\left(\frac{\mu}{t} \right)^\alpha}.$$

This simplifies to

$$h(t) = \frac{\alpha}{t}.$$

Now we will use the method of maximum likelihood to estimate $\hat{\alpha}$

$\square$

Case Studies

For each of the following case studies, create a structured abstract no longer than 4 pages in length (including figures, tables, and references). The Background section is provided for each and should be included in your write-up. You must write the Methods, Results, and Conclusion sections. Code should be included in an appendix as well.

1. The first Case Study looks at bladder cancer, both its recurrence and death due to bladder cancer. The data can be found in the file `bladder.txt`. Variables include subject's `id`, `time` (to recurrence, death, or censored in months), first recurrence status (`status1`, coded 1 for first recurrence, 0 for censored or dead), death status (`status2`, coded 1 for dead, 0 for censored or first recurrence), `treatment` (code `placebo`, `pyridoxine`, and `thiotepa`), the `number` of initial `number` tumors (coded `One`, `Two or Three`, and `Four or More`), and the `size` of the largest tumor (coded `1cm`, `2cm to 3cm`, and `4cm or Larger`).

   ## Background

   Bladder cancer is a common type of cancer, typically beginning in the urothelial cells of the bladder. The cancer is often caught in the early stages and is easily treatable. However, even early-stage surgical interventions may not prevent recurrence. To evaluate the effectiveness of post-surgical treatments, we conducted a three arm study comparing the effects of treatment with pyridoxine, thiotepa, or placebo on the primary end point of recurrence. We also consider the secondary end point of death. The size of the initial tumor or the number of initial tumors may also impact both study endpoints. We enrolled 118 patients with confirmed bladder cancer whose tumors were surgically removed and followed them until recurrence, death, or both.

2. The second Case Study examines the time to first and second recurrence of chronic granulomatous disease or CGD. As only a subset of the patients had the second recurrence, there are two datasets for this study: `cgd1.txt` (first recurrence data) and `cgd2.txt` (second recurrence data). The variables in both datasets have the same names. Variables include subejct's `id`, the `time` to recurrence of infection in days, `status` (1 for recurrence, 0 for censored), `treatment` (coded `rIFN-g` for gamma r-interferon, `placebo`), pattern of inheritance (`inherit`, coded `autosomal` or `X-linked`), and use of prophylactic antibiotics at study entry (`propylac`, coded 1 for yes, 0 for no).

   ## Background

   Chronic granulomatous disease or CGD is a genetic disorder in which white blood cells called phagocytes are unable to kill certain types of bacteria and fungi. Consequently, patients suffering from CGD are subject to frequent and potentially life-threatening infections. CGD impacts the effectiveness of phagocytes, a type of white blood cell, that makes patients more susceptible to bacterial and fungal infections. We conducted a placebo controlled trial of a cytokine-based treatment using interferon gamma to examine its effect on recurrence of CGD. The primary endpoint of the study was the first recurrence of a bacterial or fungal infection and the secondary endpoint was the second recurrence of infection. The pattern of inheritance and use of antibiotics may impact recurrence as well. In total, 128 patients were available for analysis of the primary endpoint while only 44 patients were available for the secondary endpoint.