Exponential Model

James Yang

March 31, 2022

1 Introduction

The exponential model is one where each arm i is assumed to follow an exponential distribution with hazard λ_i. Depending on the design procedure (the test statistic) there are numerous choices of grid spaces and parametrizations of the natural parameter space. This document will focus on the log-rank statistic with two arms (control and a treatment).

2 Model Assumptions

Assume that there are n patients in each of the $d = 2$ arms with independent $X_{ci} \sim E(\lambda_c), X_{ti} \sim E(\lambda_t)$, $i = 1, \ldots, n$. $X_{c\cdot}$ are the samples for the control arm and $X_{t\cdot}$ are for the treatment arm. Then, the distribution of $X \in \mathbb{R}^{n\times 2}$ forms an exponential family with sufficient statistic $T(x) = \left(\sum\limits_{i=1}^n x_{ci}, \sum\limits_{i=1}^n x_{ti}\right)$ natural parameter $\eta = (-\lambda_c, -\lambda_t)$, and log-partition function $A(\eta) := -n\log(\eta_c\eta_t)$.

3 Grid Space

Since the log-rank statistic only depends on the hazard rate $h = \lambda_t/\lambda_c$, it is convenient to parametrize the natural parameter space as a function of (λ_c, h). Moreover, we will see in Section 4.4 that we get major computation benefits of parametrizing in the log-space $\theta = (\log(\lambda_c), \log(h))$.

This parametrization defines a mapping $\eta(\theta) = \left(-e^{\theta_1}, -e^{\theta_1+\theta_2}\right)$ from the grid space to the natural parameter space. We conclude this section with the Jacobian and hessian computations needed in the later sections.

$$D_\theta \eta(\theta) = \begin{bmatrix} -e^{\theta_1} & 0 \\ -e^{\theta_1+\theta_2} & -e^{\theta_1+\theta_2} \end{bmatrix} \tag{1}$$

$$\nabla^2_\theta \eta_1(\theta) = -e^{\theta_1} e_1 e_1^\top \tag{2}$$

$$\nabla^2_\theta \eta_2(\theta) = -e^{\theta_1+\theta_2} \vec{1}\vec{1}^\top \tag{3}$$

where e_i is the ith standard basis vector and $\vec{1}$ is a vector of ones.

4 Upper Bound

For any model, we must be able to compute the upper bound estimate. The generalized upper bound estimate requires model-specific quantities, which are given by

$$\begin{aligned}
\text{Gradient Term} &: T(x) - \nabla_\eta A(\eta) \\
\eta \text{ transform} &: D_\theta \eta(\theta) v \\
\text{Covariance quadratic form} &: u^\top \operatorname{Var}(T)_\eta \, u \\
\text{Max covariance quadratic form} &: \sup_{\theta \in R} \left[v^\top (D\eta(\theta))^\top \operatorname{Var}(T)_{\eta(\theta)} (D\eta(\theta)) v \right] \\
\text{Max covariance and } \eta \text{ hessian} &: ||v||^2 \sum_{k=1}^{d} \sup_{\theta \in R} \left[||\nabla^2 \eta_k(\theta)||_{op} \sqrt{\operatorname{Var}(T_k)_{\eta(\theta)}} \right]
\end{aligned}$$

for any $v, u \in \mathbb{R}^d$ and a bounded subset $R \subseteq \mathbb{R}^d$.

The next few subsections will derive the formulas for each of the quantities above.

4.1 Gradient Term

As shown in Section 2, we have the form for $T(x)$ and $A(\eta)$.

$$\nabla_\eta A(\eta) = -n \left(\eta_c^{-1}, \eta_t^{-1} \right) = n \left(\lambda_c^{-1}, \lambda_t^{-1} \right)$$

This gives us

$$T(x) - \nabla_\eta A(\eta) = \left(\sum_{i=1}^{n} x_{ci} - n\lambda_c^{-1}, \sum_{i=1}^{n} x_{ti} - n\lambda_t^{-1} \right)$$

4.2 η Transform

Using Eq. 1, for any $v \in \mathbb{R}^d$,

$$D_\theta \eta(\theta) v = - \begin{bmatrix} e^{\theta_1} & 0 \\ e^{\theta_1 + \theta_2} & e^{\theta_1 + \theta_2} \end{bmatrix} v = - \begin{bmatrix} e^{\theta_1} v_1 \\ e^{\theta_1 + \theta_2} (v_1 + v_2) \end{bmatrix} = - \begin{bmatrix} \lambda_c v_1 \\ \lambda_t (v_1 + v_2) \end{bmatrix}$$

4.3 Covariance Quadratic Form

The covariance of T is given by

$$\operatorname{Var}(T)_\eta = n \begin{bmatrix} \eta_c^{-2} & 0 \\ 0 & \eta_t^{-2} \end{bmatrix} = n \begin{bmatrix} \lambda_c^{-2} & 0 \\ 0 & \lambda_t^{-2} \end{bmatrix} \tag{4}$$

and so,

$$u^\top \operatorname{Var}(T)_\eta \, u = n(u_1^2 \lambda_c^{-2} + u_2^2 \lambda_t^{-2})$$

4.4 Max Covariance Quadratic Form

Using Eq. 1, 4,

$$\begin{aligned}
D_\theta \eta(\theta)^\top \operatorname{Var}(T)_\eta \, D_\theta \eta(\theta) &= n D_\theta \eta(\theta)^\top \begin{bmatrix} \eta_c^{-2} & 0 \\ 0 & \eta_t^{-2} \end{bmatrix} \begin{bmatrix} e^{\theta_1} & 0 \\ e^{\theta_1 + \theta_2} & e^{\theta_1 + \theta_2} \end{bmatrix} \\
&= n D_\theta \eta(\theta)^\top \begin{bmatrix} e^{-2\theta_1} & 0 \\ 0 & e^{-2(\theta_1 + \theta_2)} \end{bmatrix} \begin{bmatrix} e^{\theta_1} & 0 \\ e^{\theta_1 + \theta_2} & e^{\theta_1 + \theta_2} \end{bmatrix} \\
&= n \begin{bmatrix} e^{\theta_1} & e^{\theta_1 + \theta_2} \\ 0 & e^{\theta_1 + \theta_2} \end{bmatrix} \begin{bmatrix} e^{-\theta_1} & 0 \\ e^{-(\theta_1 + \theta_2)} & e^{-(\theta_1 + \theta_2)} \end{bmatrix} \\
&= n \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}
\end{aligned}$$

Note the incredible simplification due to our choice of the η transformation. This gives us

$$\sup_{\theta \in R} \left[v^\top D_\theta \eta(\theta)^\top \mathrm{Var}\left(T\right)_\eta D_\theta \eta(\theta) v \right] = n v^\top \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} v$$

4.5 Max Covariance and η Hessian

From Eq. 2, 3,

$$\begin{aligned} ||\nabla^2 \eta_1(\theta)||_{op} &= e^{\theta_1} ||e_1 e_1^\top||_{op} = e^{\theta_1} \\ ||\nabla^2 \eta_2(\theta)||_{op} &= e^{\theta_1 + \theta_2} ||\vec{1}\vec{1}^\top||_{op} = e^{\theta_1 + \theta_2} d \end{aligned}$$

This gives us

$$||v||^2 \sum_{k=1}^{d} \sup_{\theta \in R} \left[||\nabla^2 \eta_k(\theta)||_{op} \sqrt{\mathrm{Var}\left(T_k\right)_{\eta(\theta)}} \right] = ||v||^2 \sqrt{n}\left(1 + d\right)$$