
MATH 2625: Biostatistical Methods

Homework 6, due Tuesday, April 29

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. Recall that the Nelson-Aalen estimator of the survivor function is

$$\tilde{S}(t) = \prod_{j=1}^k \exp\left(-\frac{d_j}{n_j}\right).$$

Using a similar “proof” to what we used for the Kaplan-Meier standard error estimate, show that the standard error of the Nelson-Aalen estimator is

$$SE\{\tilde{S}(t)\} \approx \tilde{S}(t) \left\{ \sum_{j=1}^k \frac{d_j}{n_j^2} \right\}^{1/2}.$$

This will require using the Poisson approximation to the Binomial distribution, specifically if $D_j \sim \text{Binom}(n_j, p_j)$ then $D_j \dot{\sim} \text{Pois}(n_j p_j)$ for sufficiently large n_j and small p_j .

Proof. The Nelson-Aalen survival function estimator is defined as

$$\tilde{S}(t) = \prod_{j=1}^k \exp\left(-\frac{d_j}{n_j}\right),$$

where n_j is the total number of subjects at risk at time t_j and d_j is the number of events at time t_j . When we take the natural logarithm of $\tilde{S}(t)$ and simplify, we get

$$\log \tilde{S}(t) = - \sum_{j=1}^k \frac{d_j}{n_j}.$$

Let $\hat{p}_j = \frac{d_j}{n_j}$. Then

$$\log \tilde{S}(t) = \sum -\hat{p}_j.$$

Let $D_j \sim \text{Bin}(n_j, p_j)$. Then d_j is a realization of D_j . For sufficiently large n_j and small p_j , we can use the poisson approximation of the binomial distribution, and observe that $D_j \sim \text{Pois}(n_j p_j)$. This means $\text{Var}(d_j) \approx n_j p_j$. Then

$$\text{Var}(\hat{p}_j) \approx \left(\frac{1}{n_j}\right)^2 \text{Var}(d_j) = \frac{d_j}{n_j^2}.$$

Assuming event times are independent,

$$\text{Var}(\log \tilde{S}(t)) \approx \sum \frac{d_j}{n_j^2}.$$

Now we will examine the natural logarithm of the survivor function. Using the delta method, we can see that asymptotically for large n ,

$$\text{Var}(\log \tilde{S}(t)) \approx \frac{1}{\tilde{S}(t)^2} \text{Var}(\tilde{S}(t)).$$

Multiplying both sides by $\tilde{S}(t)$, we get

$$\tilde{S}(t)^2 \text{Var}(\log \tilde{S}(t)) \approx \text{Var}(\tilde{S}(t)).$$

Substituting in what we found for $\text{Var}(\log \tilde{S}(t))$ earlier, we get

$$\text{Var}(\tilde{S}(t)) \approx \tilde{S}(t)^2 \sum \frac{d_j}{n_j^2}.$$

To get the standard error, we take the square root of the variance, leaving us with

$$SE \{ \tilde{S}(t) \} \approx \tilde{S}(t) \sqrt{\sum \frac{d_j}{n_j^2}}.$$

□

Case Studies

For each of the following case studies, create a structured proposal no longer than 4 pages in length (including figures, tables, and references). Use the accompanying articles to construct your own background section—you may base the background on the information in the papers, but it must be in your own words. Then write a methods section that includes at least a description your proposed study design, your hypothesized effect size with support, your power and sample size analysis (i.e. how many subjects are required to attain what power for the hypothesized effect size), the methods you will use to analyze the data once it is collected, and how long you plan to run the study. Your discussion section may be brief but should detail potential limitations of the study and any ethical issues that may arise. Code should be included in an appendix as well. Both case studies involve recently published reports on ongoing clinical trials examining an RSV vaccine. While there is overlap in the topics, the backgrounds of each case study should be written separately since the target populations differ. You will be proposing a replication study based on their interim results. Be sure to cite all relevant statements and support, including citations to the motivating papers.

1. Using Kampmann et al. (2023), alongside this assignment, design a replication study to assess the bivalent prefusion F vaccine in pregnant women. The goal of the study is to prevent RSV illness in infants and mothers postpartum.
2. Walsh et al. (2023), also accompanying this assignment, discusses the efficacy and safety of a bivalent RSV prefusion F vaccine in older adults. Using it as a guide, design a replication study to assess the efficacy and safety of the vaccine in adults aged 60 years or older.