

*BIOS 522: Survival Analysis Methods*

**Activity 2:**

**The survival function**

*This week, we defined the survival function for time-to-event data. We learned about two non-parametric methods for estimating the survival function: the empirical CDF when censoring is absent, and the Kaplan-Meier method when censoring is present.*

Problem 1. Glioma data

The following data summarize weeks to death or censoring in 10 adults with recurrent astrocytoma (a type of cancer that occurs in the brain or spinal cord).

|  |
| --- |
| Astrocytoma: 6, 6+, 13, 21, 30, 31+, 37, 38, 47+, 49 |

1. Calculate the Kaplan-Meier estimate of the survival function. Though you may use a calculator, show your work to demonstrate how it is done by hand. Report the entire function in intervals of time. Only report intervals where the survival function changes.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **# at risk** | **# of deaths** | **# censored** | **Conditional survival probability** | **Kaplan-Meier estimate** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

The final estimate should collapse across intervals where there is no change, i.e., and

1. Report a point estimate and standard error for survival at time 21 weeks.

Point estimate at 21 weeks drops to 0.675.

Greenwood’s variance estimator is:

And the standard error is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **# at risk** | **# of deaths** | **# censored** | **Variance term** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Problem 2. Discussion board post

Have each small group participant take 1-2 minutes to describe the Kaplan-Meier curve example they posted in this week’s discussion board. As a group, identify one example to share with the class. Prepare a new discussion board post that summarizes the example, making sure to include the elements listed below. Select a group member to present this to the class. (Ideally this will *not* be the person who originally identified the plot, though that person can offer comments if there is subsequent discussion.)

1. *Label the post with your group number*
2. *Provide a reference/link to the study*
3. Goal: [In one or two sentences, describe the primary scientific goal that the study investigators sought to address.]
4. Population: [In one or two sentences, describe the population included in the study. Examples you might include (if relevant to the study): sample size, age, location, time window, and eligibility criteria.]
5. Outcome variable: [In one or two sentences, describe the time-to-event variable being studied. This should include the time origin and the definition of the event. Optional - describe any censoring.]
6. Predictor variables: [If more than one Kaplan-Meier curve is plotted, describe the categorical grouping used (e.g. age >40, <=40 years.)]
7. Results: [Provide a copy of the figure. In two to three sentences, describe the main findings. It is not necessary to address questions of statistical significance.]
8. What do you notice? List four things you notice about the plot.