

*BIOS 522: Survival Analysis Methods*

**Homework - Weeks 4-5**

*Homework assignments contain a mix of theoretical, applied, and computing questions. Students are encouraged to work with classmates but must prepare their answers individually. Copying and pasting between students is not allowed. Assignments are graded on presentation, and points will be deducted from messy or illegible answers.*

Problem 1. Exponential maximum likelihood estimator (1.5 points)

Recall that the likelihood function for the right-censored data assuming an exponential distribution is as follows:

Demonstrate that the maximum likelihood estimator of is the incidence rate:

*Hint: Calculate the score function (the first derivative of the log likelihood function) and set this equal to zero. Solve for .*

Set equal to zero and solve for  :

Problem 2. The hazard function across scales (1.5 points)

Like incidence rates, the hazard function has lower limit 0 and no upper limit (i.e., it can exceed one). The hazard function has units of 1/time (time-1).

Consider an exponential random variable . Imagine that the mean survival time is 2 days.

1. (0.5 points) What is the corresponding hazard rate? Include units.

The corresponding hazard is 0.5 days-1.

1. (0.5 points) What if we described the mean survival time of as 48 hours instead of 2 days. What is the corresponding hazard rate? Include units.

The corresponding hazard is or 0.021 hours-1.

1. (0.5 points) What if we described the mean survival time of in weeks instead of hours or days. What is the corresponding hazard rate? Include units.

The mean survival time is 2/7 or 0.286 weeks. The corresponding hazard is 3.5 weeks-1.

Problem 3. More glioma data (3 points)

The following data summarize weeks to death or censoring in 10 adults with recurrent astrocytoma gliomas.

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

1. (1 point) Use the astrocytoma data to calculate the Nelson-Aalen estimate of the cumulative hazard function at weeks. Show your work.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Unique failure/ censoring time** | **Number at risk during** | **Number of deaths at** | **Number censored at** | **Cumulative hazard contribution** | **Nelson-Aalen estimate** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

0.368 or 36.8%

1. (0.5 points) Report an accompanying standard error for the cumulative hazard function at time 25 weeks.

Standard error = 0.2146

1. (0.5 points) Report a 95% log-transformed confidence interval for the cumulative hazard function at time 25 weeks.
2. (0.5 points) Report a point estimate of survival at time 25 weeks calculated by the Breslow estimator.
3. (0.5 points) Report a 95% complementary log-log transformed confidence interval for survival at 25 weeks.

Or transform the interval from part c.

Problem 4. Working with likelihoods (2 points)

For this question, we will continue to use the data from 10 adults with recurrent astrocytoma glioma.

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

Imagine that we fit a Weibull distribution to the data and return maximum likelihood parameter estimates   and  . Write out the likelihood contribution for the first two individuals, evaluated at the maximum likelihood estimates   and .

For the first individual, :

For the second individual, :

Problem 5. Order matters! (2 points)

Consider Cox’s partial likelihood:

An interesting fact is that an analysis based on Cox’s partial likelihood is invariant to any rank-preserving transformation of the follow-up times . The *times* do not matter in the above; what matters is the *order* in which the failure/censoring times occur. We could change the times, but as long as we preserve the order, we will make the same inference.

In 2-4 sentences, explain to your classmate why this is true from the form of the likelihood.

The times themselves do not appear in the partial likelihood. is a vector of baseline covariates, so the terms are not dependent upon time. For each unique failure time, as long as the risk sets are the same, the inference will be the same. The risk sets will be the same under a rank-preserving transformation.