

*BIOS 522: Survival Analysis Methods*

**Activity 6:**

**Interpreting the Cox model**

*This week, we interpreted hazard ratios for binary, categorical, and continuous covariates. We extended the allowable shapes by transformations, splines, and interactions. We defined global and local hypothesis tests for the Cox model and constructed confidence intervals for hazard ratios.*

Problem 1. Calculating hazard ratios

Using data from the ACTG 320 trial described in the Week 3 R Tutorial, we fit the following Cox proportional hazards model:

where if the participant receives indinavir (0 otherwise), is the participant’s baseline Karnofsky score, and is the participant’s baseline CD4 cell count (cells/mm3). Karnofsky score measures health and functional impairment, with patients having scores of 70 (most impairment), 80, 90, or 100 (no impairment). The outcome is new AIDS-defining event or death.

We fit the model with partial likelihood. We obtain the following coefficients:

|  |  |  |
| --- | --- | --- |
| **Covariate** | **Coefficient** | **Exp(Coefficient)** |
| Treatment |  |  |
| Karnofsky score = 80 |  |  |
| Karnofsky score = 90 |  |  |
| Karnofsky score = 100 |  |  |
| CD4 |  |  |

Using the coefficients in the table above, calculate the estimated hazard ratio comparing each of the following sets of populations.

1. Participants with tx = 1, Karnofsky score = 100, CD4 = 50 *vs.*

Participants with tx = 1, Karnofsky score = 70, CD4 = 50.

There are several ways to approach this problem. We could start from first principles (the long way):

We could plug in the coefficients at this point and calculate the hazard ratio.

Alternatively, we can see that more of these terms will cancel if we keep going.

Thus, the hazard ratio for a population with Karnofsky = 100 versus a population with Karnofsky = 70 (as long as everything else is the same) is captured by a single coefficient,  .

The short way to answer this problem would be to recognize that since only Karnofsky score has changed, the answer is:

1. Participants with tx = 1, Karnofsky score = 70, CD4 = 70 *vs.*

Participants with tx = 1, Karnofsky score = 70, CD4 = 50.

Again, there are several ways to approach this problem. We will start from first principles (the long way):

We could plug in the coefficients at this point and calculate the hazard ratio.

Alternatively, we can see that more of these terms will cancel if we keep going.

If we know our rules for exponentials, this simplifies even more:

Thus, since the coefficient for CD4 reflects a one unit-change in CD4, then we must multiply this coefficient by 20 to calculate a twenty-unit change in CD4.

Another way to approach the problem – a one-unit change is:

A two-unit change is:

A twenty unit change is:

1. Participants with tx = 1, Karnofsky score = 100, CD4 = 50 *vs.*

Participants with tx = 1, Karnofsky score = 80, CD4 = 50.

Again, there are several ways to approach this problem. We will start from first principles (the long way):

We could plug in the coefficients at this point and calculate the hazard ratio.

Alternatively, we can see that more of these terms will cancel if we keep going.

Thus, the hazard ratio comparing two non-reference groups can be calculated from the ratio of the hazard ratios (comparing each to the reference).

Thus,

Another way to see this is the following (using informal notation):

1. Participants with tx = 1, Karnofsky score = 90, CD4 = 50 *vs.*

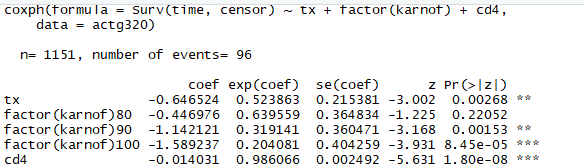
Participants with tx = 0, Karnofsky score = 70, CD4 = 50.

Now both treatment and Karnofsky score have changed, so we need to approach the problem from first principles. Let’s estimate the hazard for each group and then take the ratio.

So we can plug in our numbers to the above, or we can simplify further.

Thus, we can calculate the hazard ratio reflecting these multiple changes by multiplying the hazard ratios for each effect:

R code is provided to fit the Cox model for the ACTG320 data set.



Problem 2. Calculating test statistics

Consider the leukemia data set in the R survival package. The dataset summarizes survival in 23 patients with Acute Myelogeneous Leukemia. The clinical question was whether the standard course of chemotherapy should be extended (“maintenance”) for additional cycles.

We are interested in testing the null hypothesis that maintenance has no impact on survival. We fit two models:

* A model with no covariates. It has log likelihood -.
* A model with a single binary covariate for maintenance. It has MPLE  , variance of   , and log likelihood -41.03

1. Calculate a Wald test statistic for .

0.916^2/0.262 = 3.206

1. Calculate a likelihood ratio test statistic for .

-2(-42.73+41.03) = 3.4

1. For this model with only a single binary covariate, describe how we might calculate a score test statistic for .

Calculate a log-rank test

See code below

