Applying Cox Regression to Competing Risks

Hanna Damarjian

11/18/2021

Introduction.

When it comes to Survival Analysis, one of the most vital questions to inquire about a patient's death is the failure (risk) or *specific cause* of death. When the time of a diagnosis and failure time are known, the measured survival time becomes known information (a condition), and then we can use it to predict the *rate* of other patients dying. This is where Cox Regression modelling can be applied where we are focusing on a single covariate, x, and single failure type, δ. However, what happens if a patient has more than diagnosis or failure type? The answer to this question becomes more difficult because these types of failures may be difficult to discern to the specific cause of death. A proposed idea is to use Cox Regression modelling, but how can we build this model when a patient is known to have at least two or more risk factors? It is analogous to thinking that these risk factors are **competing** to be the leading cause of death. In other words, if we do not necessarily know the type of failure (that caused a patient's death), can we adequately model this rate regardless of knowing the leading type of failure? This paper develops two methods using Cox Regression modelling and shows that it is possible to analyze competing risks using readily available standard programs. Each of the two methods assumes *two types of failures* and some censored data. I discover that the first method approaches finding the partial likelihood for each failure type (denote these events as A and B) from $P(A \cap B) = P(A) + P(B) - P(A \cup B)$, and the second approaches finding the partial likelihood for each failure type through stratificiation. I thought of stratification as $P(A \cap B) = P(A)P(B)$ where events A and B are independent. Applications are illustrated, but much of this paper will focus on my explanation and comprehension of the theoretical methodology.

Data Set-Up for both methods.

Assumptions that were made for both methods include:

1. $\forall i = 1, 2, ..., m$ indivdiuals, there are two failure types: δ_i and $1 - \delta_i$ where $\delta_i = 1$ if the ith patient died due to failure type 1 (implying $(1 - \delta_i) = 0$ for failure type 2) and vice versa. Note that the paper does not consider the interaction (or joint relationship) between failure types.

2. Data for response failure time, t_i, is duplicated twice. If an individual has failure type 1 (for example), then failure type 2 is censored, *, and vice versa.

3. If the subject is censored (i.e. t_i^*), then we make two entries as explained above, however, the duplicated failure time t_i will be censored for **both** failure types.

4. If subject i has failure time, t_i and $\delta_i = 0$, then the covariates include $x_i, (\delta_i = 0)x_i = 0$. As a pair, the paper specifies this as (**x,0**). Analogously, if $\delta_i = 1$, then the covariate pair would be (**x,x**). Thus, the Cox regression model for the ith individual under failure type 1 would be affected by covariates $x_i, \delta_i x_i$. We would then regress on the the duplicated covariates.

5. The hazard functions for the two types of risks are assumed to be additive. That is, the hazard of failure 1 or 2 is $x_i + \delta_i x_i$.

6. A failure type is specified as the cause of death by taking $\mathbf{min}\{t_{i;failuretype_1}, t_{i;failuretype_2}\}$. For example, if $t_{i;failuretype_1}$ is the smaller value between the two failure times, then we use x_i as the covariate part of our model.

7. When observation ceases due to failure or censoring, two survival times of the same duration have been observed, one for each process (model). At least one of the times being censored and the interaction of x_i with failure type is also considered.

Method 1.

Theory Part A presented:

Assuming no ties, the contribution to the partial likelihood if observation i results in a failure (type 1 or type 2) is:

$$L(b_0, b, \theta) = \frac{exp(b_0\delta_i + b'x_i + \theta'\delta_i x_i)}{\sum_{R_i} exp(b_0\delta_i + b'x_i + \theta'\delta_i x_i)} \tag{1}$$

where θ is a parameter that represents the difference between two vectors of regression coefficients for the failure types. This is interpreted as *"the probability that subject i has failure type δ given all those subjects available for failure of either type."* The contribution to the denominator from the two duplicated entries with original covariate, x is $\exp(b'x)$ which comes from the covariate pair (**x,0**) and $\exp(b_0 + b'x + \theta'x)$ which comes from (**x,x**).

This is all derived under the assumption that the hazard functions for the two failure types for an individual with covariate x are proportional to $\exp(b'x)$ and $\exp(b_0 + b'x + \theta'x)$ with some baseline hazard function $\lambda_{01}(t)$.

How is Theory Part A mathematically implemented?

Recall that if we want to compare two groups, then we use the partial likelihood function with: $h_1(t_i) = e^{\beta x_i} h_0(t_i)$. Also under the assumption that no ties occurred for failure times, Cox shows that

$$L(\beta) = \prod_{i=1}^{m} \frac{exp(\beta x_i)}{\sum\limits_{k \in R_{t_i}} exp(\beta x_k)} \tag{2}$$

where $R(t_i)$ is the risk set of subjects at t_i. We want to only estimate (β) and not (β, t_i). Hence the name *partial likelihood*. The idea is very similar here, except that instead of βx_i as part of the defined rate function, $r(x, \beta)$, we have the additive form $b_0 \delta_i + b' x_i + \theta' \delta_i x_i$, where we want to estimate the likelihood of b_0, b, and θ'. Note that b_0 represents the regression coefficient for the failure type (for every ith subject), b is the regression coefficient for the covariate, and θ' is the regression coefficient mentioned prior. The hazard function set up for $\lambda_{01}(t)$ and $\lambda_{02}(t)$ are derived and defined from the hazard function having the general form $h_i(t_i, x_i, \beta) = h_0(t_i) r(x_i, \beta)$. For example, the hazard function for failure type 1 of a subject with covariate x is $e^{bx} \lambda_{01}(t)$ whereas the the hazard function for failure type 2 of a subject with covariate x is $e^{b_0 + bx + \theta' x} \lambda_{02}(t)$.

Theory Part B presented

The paper then focuses on finding the parameter estimates for $b := b_I$ and $b + \theta' := b_{II}$ which are the vector coefficients appropriate to each failure type. Assuming no ties, the full partial log-likelihood is:

$$L(b_I, b_{II}) = \sum_{j,I} b_I' x_j + \sum_{j,II} (b_0 + b_{II}' x_j) - \sum_{j} ln[\sum_{R_j} exp(b_I') + exp(b_0 + b_{II}' x)]) \tag{3}$$

where $\sum\limits_{j,I} b_I' x_j$ is the sum of all type 1 failures, $\sum\limits_{j,II} (b_0 + b_{II}' x_j)$ is over the type II failures, and $\sum\limits_{j} ln[\sum\limits_{R_j} exp(b_1') + exp(b_0 + b_{II}' x)])$ is over all failures (type 1 or type 2).

Theory Part B mathematically implemented?

Denote X to be a collection of random variables that takes discrete or continuous values based upon $i = 1, 2, ..., m$ individuals. Thus, $X = (X_1, X_2, ... X_m)$ and X_i represents the arbitrary covariate for the jth individual.

Now, let U be the universal set of two types of failures - A if failure type 1 is the leading cause $(\delta_i = 0)$

and B if failure type 2 is the leading cause ($\delta_i = 1$) - such that $A, B \subset U$. According to set theory and from what we have defined, there are only four types of events that can occur: $A, B, A \cap B$, and $A \cup B$. The complement $A \cup B$ does not exist because U is defined to consist of only two failure types (there does not exist any other failure type).

Suppose each event's measure denotes the hazard ratio between one of the four events at $X_i = x_i$ and (assuming) the baseline is $X_i = 0$ for the ith patient. For example, the hazard ratio for patient $i = 15$ (given $m = 20$) under event A with covariate age $X_{15} = 49$ is $\frac{e^{b_0(\delta_{15}=0)+b_I(x_{15}=49)+b_{II}(\delta_{15}=0)(x_{15}=49)}}{e^{b_0(\delta_{15}=0)+b_I(x_{15}=0)+b_{II}(\delta_{15}=0)(x_{15}=0)}} = \frac{e^{b_0(\delta_{15}=0)+b_I(x_{15}=49)+b_{II}(\delta_{15}=0)(x_{15}=49)}}{1} = e^{b_I*(49)}$, where b_0 is the unknown parameter estimate for the intercept, b_I is the parameter estimate for covariate $x_{i=15}$ (age for 15th patient) and b_{II} is the interaction parameter estimate between failure type, $\delta_{i=15}$ and x_{15}. The hazard ratios (measures) for events B and $A \cup B$ (given ith individual and covariate x_i) are $e^{b_0+b_{II}x_i}$ and $\sum_{R_i}(e^{b_I*x_i} + e^{b_0+b_{II}x_i})$ where R_i is the risk set of the ith individual. This occurs due to failure type censoring. We defined that if the patient dies from failure type 1 (as its leading cause) then the hazard ratio is $e^{b_I*x_i}$. But $e^{b_I*x_i} + e^{b_0+b_{II}x_i}$ denotes a patient dying from one or the other failure type (not both).

Next, $A \cap B$ cannot be defined because, mathematically speaking, δ_i is one of two outcomes and not both 0 and 1 ($0 \neq 1$). Pragmatically speaking, however, it *can* occur in which the ith patient with x_i can die because of both failure types 1 and 2.

Recall that the goal is to find $L(b_0, b_I, b_{II}) := L$. How can we find L using A, B, and $A \cup B$? We do not have $A \cap B$, but we can use the Additive Theorem of Probability to derive $P(A \cap B)$ using $P(A), P(B)$, and $P(A \cup B)$ and given (practically speaking), some data set with sample points $(x_i, \delta_i x_i)$, we can then find L and solve for parameter estimates b_0, b_I, b_{II}. In other words, we are building a function for $L(b_0, b_I, b_{II}; A \cap B)$ using $L(b_0, b_I, b_{II}; A), L(b_0, b_I, b_{II}; B)$, and $L(b_0, b_I, b_{II}; A \cup B)$.

From above, we first use the Additive Theorem of Probability:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A) + P(B) - P(A \cup B),$$

rewritten as:

$$P(A \cap B; b_0, b_I, b_{II}) = P(A; b_0, b_I, b_{II}) + P(B; b_0, b_I, b_{II}) - P(A \cup B; b_0, b_I, b_{II})$$

Using the given information, the following probabilities are computed as:

$$P(A;b_0, b_I, b_{II}) = \frac{A}{U} = \frac{A}{(A \cup B)} = \frac{e^{b_I x}}{e^{b_1 x} + e^{b_0 + b_{II} x}},$$

$$P(B; b_0, b_I, b_{II}) = \frac{B}{U} = \frac{B}{(A \cup B)} = \frac{e^{b_0 + b_{II} x}}{e^{b_1 x} + e^{b_0 + b_{II} x}},$$

$$P(U; b_0, b_I, b_{II}) = P(A \cup B; b_0, b_I, b_{II}) = \frac{A \cup B}{U} = \frac{(A \cup B)}{(A \cup B)} = \frac{e^{b_1} + e^{b_0 + b_{II} x}}{e^{b_1} + e^{b_0 + b_{II} x}}.$$

I will not simplify $P(A \cup B)$ as $P(A \cup B) = 1$ for further analysis below. Hence, we have:

$$\begin{aligned} P(A \cap B; b_0, b_I, b_{II}) =& P(A; b_0, b_I, b_{II}) + P(B; b_0, b_I, b_{II}) - P(A \cup B; b_0, b_I, b_{II}) \\ =& \frac{e^{b_I x}}{e^{b_I} x + e^{b_0 + b_{II} x}} + \frac{e^{b_0 + b_{II} x}}{e^{b_I x_i} + e^{b_0 + b_{II} x}} - \frac{e^{b_I} x + e^{b_0 + b_{II} x}}{e^{b_I} x + e^{b_0 + b_{II} x}}. \end{aligned}$$

Since the parameter estimates are unknown, we want to maximize $b_0, b_I.b_{II}$ if we have given sample points $(x_i, \delta_i x_i)$ that meets the conditions from above. Finding the data is not difficult, but we are more interested in finding $b_0, b_I.b_{II}$ that *maximize* the probability of our sample producing those sample points. So, for example of event A, we want to maximize the following function:

$$L(b_0, b_I, b_{II}; A) = P(A; b_0, b_I, b_{II}).$$

Note that this is for one event A. Since we have $i = 1, 2, ..., m$ sample points, then we have a collection of sample events, where $A = (A_1, A_2, ..., A_m)$. Also, we assumed that we have m independent data points drawn from the same probability distribution, so the joint probability distribution of $P(A; b_0, b_I, b_{II})$ is:

$$
\begin{aligned}
L(b_0, b_I, b_{II}; A) =& P(A; b_0, b_I, b_{II}) \\
=& P(A_1; b_0, b_I, b_{II}) P(A_2; b_0, b_I, b_{II}) ... (P(A_m; b_0, b_I, b_{II}) \\
=& \prod_{i=1}^{m} P(A_i; b_0, b_I, b_{II}) \\
=& \prod_{i=1}^{m} \frac{e^{b_I x_i}}{\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})},
\end{aligned}
$$

where R_i is the risk set for the ith patient (according to what Cox proposed). Doing the same for $P(B; b_0, b_I, b_{II})$ and $P(A \cup B; b_0, b_I, b_{II})$, yields:

$$
\begin{aligned}
L(b_0, b_I, b_{II}; A \cap B) =& L(b_0, b_I, b_{II}; A) + L(b_0, b_I, b_{II}; B) + L(b_0, b_I, b_{II}; A \cup B) \\
=& \prod_{i=1}^{m} \frac{e^{b_I x_i}}{\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})} + \prod_{i=1}^{m} \frac{e^{b_0 + b_{II} x_i}}{\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})} - \prod_{i=1}^{m} \frac{\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})}{\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})}. \\
=& \frac{\prod_{i=1}^{m} e^{b_I x_i}}{\prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i}))} + \frac{\prod_{i=1}^{m} e^{b_0 + b_{II} x_i}}{\prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i}))} - \frac{\prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i}))}{\prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i}))}.
\end{aligned}
$$

The denominator of $L(b_0, b_I, b_{II}; A \cap B)$ is the same as the other probabilities. Thus, if we multiply out the denomiator (which is $\prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_I} + e^{b_0 + b_{II} x_i})))$, then we have

$$
L = \prod_{i=1}^{m} e^{b_I x_i} + \prod_{i=1}^{m} e^{b_0 + b_{II} x_i} - \prod_{i=1}^{m} (\sum_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})) \tag{4}
$$

where $L := \prod_{i=1}^{m} (\sum\limits_{R_i} (e^{b_1 x_i} + e^{b_0 + b_{II} x_i})) * L(b_0, b_I, b_{II}; A \cap B)$ This makes differentiation easier when finding the parameter estimates for b_0, b_I, and b_{II}.

Finally, we can apply the ln() function to each *term* (event) on both sides of the equation. We cannot do it to each *side* of (4) because we would treat events A, B, and $A \cup B$ as one event which would contradict our definitions for each event. It also makes to take ln() for each term because what if a researcher is solely interested in the likelihood for failure type 1 (A)? The researcher would want to focus on differentiating $\prod_{i=1}^{m} e^{b_I x_i}$ and not $\prod_{i=1}^{m}(\sum\limits_{R_i}(e^{b_1 x_i} + e^{b_0 + b_{II} x_i}))$. Thus:

$$
\begin{aligned}
ln(L) =& ln(\prod_{j,I} e^{b_I' x_j}) + ln(\prod_{j,II} e^{(b_0 + b_{II}' x_j)}) - ln(\prod_{j,I,II}(\sum_{R_j} e^{b_I'} + e^{b_0 + b_{II}' x_j}))\\
=& ln(e^{\sum\limits_{j,I} b_I' x_j}) + ln(e^{\sum\limits_{j,II}(b_0 + b_{II}' x_j)}) - \sum_{j,I,II} ln(\sum_{R_j}(e^{b_I'} + e^{b_0 + b_{II}' x_j}))\\
=& \sum_{j,I} b_I' x_j + \sum_{j,II}(b_0 + b_{II}' x_j) - \sum_{j,I,II} ln(\sum_{R_j}(e^{b_I'} + e^{b_0 + b_{II}' x_j})).
\end{aligned}
$$

The paper then addresses how to find the parameter estimates by finding $\frac{\partial L}{\partial b_I}, \frac{\partial L}{\partial b_{II}}$, and $\frac{\partial L}{\partial b_0}$ by setting them all equal to 0 and solving for each estimator.

Method 2.

Theory presented: The paper states to run a Cox regression on the covariates $\mathbf{x}$, $\delta\mathbf{x}$ stratifying by failure type, $\delta = 0$ or 1. When treating the survival times of the two types of failures separately, the partial likelihood for two failure types with regression coefficients $b, b+\theta$ and unknown baseline hazard functions, h_{00}, h_{01}, is:

$$
L(b, b+\theta) = \prod_{t_i, \delta_i = 0} \frac{exp(b' x_i)}{\sum\limits_{R_i} exp(b' x_i)} \prod_{t_i, \delta_i = 1} \frac{exp(b' x_i + \theta x_i)}{\sum\limits_{R_i} exp(b' x_i + \theta x_i)}. \tag{5}
$$

Remark: the risk set for each case consists of those subjects with the appropriate stratum identifier, $\delta = 0$ for the first and $\delta = 1$ for the second.

Theory mathematically implemented?

Denote event C to be the hazard ratio for a patient led to death by failure type 1 and event D the hazard ratio for a patient led to death by failure type 2. Given sample points $(x_i, \delta_i x_i)$ for all $i = 1, 2, ..., m$ individuals, we want to find the likelihood function corresponding to $P(C_i \cap D_i)$. From the assumptions, the likelihood function according to C is $L(b', \theta; C) = P(C; b', \theta) = \prod_{i=1}^{m} P(C_i; b', \theta) = \prod_{i=1}^{m} \frac{e^{b' x_i + \theta(\delta_i = 0) x_i}}{\sum\limits_{R_i} e^{b' x_i + \theta(\delta_i = 0) x_i}} =$

$\prod_{i=1}^{m} \frac{e^{b'x_i}}{\sum\limits_{R_i} e^{b'x_i}}$ where b' is the unknown parameter estimate for the discrete or continuous covariate x_i, θ is the unknown parameter estimate for failure type 2, $\sum\limits_{R_i}$ represents the risk set $\forall i \in [1, m]$ that are uncensored, and $\delta_i = 0$ here for failure type 1. Also, from Method 1, we had an unknown parameter estimate for b_0, but we do not have an intercept parameter here because stratification of δ_i separates the probabilities. They are independent from one another which implies that we are not interested in the probability of failure type 1 or failure type 2 occuring.

Similarly, the likelihood function according to D is $L(b', \theta; D) = P(D; b', \theta) = \prod_{i=1}^{m} P(D_i; b', \theta) = \prod_{i=1}^{m} \frac{e^{b'x_i+\theta(\delta_i=1)x_i}}{\sum\limits_{R_i} e^{b'x_i+\theta(\delta_i=1)x_i}} = \frac{e^{b'x_i+\theta x_i}}{\sum\limits_{R_i} e^{b'x_i+\theta x_i}}$. Since $C = (C_1, C_2, ..., C_m)$ and $D = (D_1, D_2, .., D_m)$ are independent (due to the stratification) $\rightarrow C_i$ and D_i are independent $\rightarrow P(C_i)$ and $P(D_i)$ are also independent. Finally, if $L(b', \theta; C \cap D) := P(C \cap D; b', \theta)$, then:

$$\begin{aligned}
P(C \cap D; b', \theta) =& P(C; b', \theta) P(D; b', \theta) \\
=& \prod_{i=1}^{m} P(C_i; b', \theta) \prod_{i=1}^{m} P(D_i; b', \theta) \\
=& \prod_{i=1}^{m} \frac{e^{b'x_i}}{\sum\limits_{R_i} e^{b'x_i}} \prod_{i=1}^{m} \frac{e^{b'x_i+\theta x_i}}{\sum\limits_{R_i} e^{b'x_i+\theta x_i}}.
\end{aligned}$$

Therefore, we have $L(b', \theta; C \cap D) = \prod_{i=1}^{m} \frac{e^{b'x_i}}{\sum\limits_{R_i} e^{b'x_i}} \prod_{i=1}^{m} \frac{e^{b'x_i+\theta x_i}}{\sum\limits_{R_i} e^{b'x_i+\theta x_i}}$. In terms of time, t_i and δ_i, this is equivalent to $L(b', \theta; C \cap D) = \prod_{t_i, \delta_i=0} \frac{e^{b'x_i}}{\sum\limits_{R_i} e^{b'x_i}} \prod_{t_i, \delta_i=1} \frac{e^{b'x_i+\theta x_i}}{\sum\limits_{R_i} e^{b'x_i+\theta x_i}}$.

Critiquing the Article's Implementation.

Stanford Heart Transplant Data (Assumptions and Critique).

Survival times of 65 patients who received heart transplants were considered. Failures are death from either rejected (type 1) or other cause (type 2). The paper first illustrates estimating the hazard function for rejection, the hazard function for other cause, and the hazard function for any cause (which is incorporating both types of failures) using Kaplan-Meier estimation. From the data, it appears that since the hazard functions (rates) are not proportional (i.e. slopes are not parallel), this data would not be a good fit to utilize either Method A nor B. This is because an assumption to implement this model (as mentioned before) requires the rates to be proportional to one another. However, the authors decided to go forth and attempt each method.

Under Method A, the authors incorporated three covariates - failure type, age at transplant, and tissue mismatch score. Interaction terms between failure type and each of the other covariates was also included. From the resulting small proportional hazard model p-values for the regresstion interaction failure type estimates, the authors confirmed a poor fit.

Under Method B, the authors stratified by failure type (as mentioned in one of the assumptions for this methodology). Each of the two covariates were included in each strata along with the respective interaction term of failure type and each covariate.

My critique on this analysis is to consider an alternative data set. It would not be of beneficial use to proceed with using these two methods if a major assumption (hazard proportional models for each failure type set-up) is most-likely violated. This was due to the cause-specific survival curves crossing. From class, we know that the hazard function depends upon the probability density function of t and the survival time. From the tables, the parameter estimates and their significance would not be of much use, especially with interpretability with this data set. I recall that in class we would focus on determining if the hazard functions were linear and parallel according to the $ln(-ln(S(t)))$ vs. $ln(t)$ graph. Violation of this assumption implies that the Cox Proportional Hazard Model would not be a good fit to the data. This data set illustrates the violation. Another concept I can connect is how we spent time concocting the partial likelihood function for a single sample (of failure) under the Cox Proportional Hazard Model set-up. The two methods discussed are very similar since we are trying to find the partial likelihood function. However, the difference is that this example illustrates two samples of failure under the CPH model set-up.

Prostate Cancer Data (Assumptions and Critique).

Survival times of 506 patients with prostate cancer were randomly allocated to different levels of treatment with thedrug diethylstilbestrol (DES). There eight specific risk factors (including drug treatment), and 23 subjects had incomplete data. The failures or causes of death were classifed as cancer, cardiovascular, or other. The resulting survival probability curve shows that two causes (cancer and cardiovascular) cross. Since competing risks data is cause-specific, and if surivival probability curves cross, then it would be difficult to compare risk factors to causes of a patient's death. That is, it would be difficult to discern which risk factor was ultimately the leading factor to a patient's death (failure). The illustration on page

7 of the paper shows that Method B would be appropriate to use and not Method A. This is because the assumption of no ties is violated and occurs when the survival probability curves of cardiovascular (CVD) and cancer cross.

However, the authors took an interesting approach by ommitting patients with failure time less than 5 months and refitting the survival probability curves. This assumption was met as the two survival curves do not cross (no ties). Hence, Method A can be utilized for CPH model fitting. On page 8, the author lays a table illustrating how a simple model between CVD and other failure types as the only covariates have insignficant p-values. I find it interesting to see that other causes of death is nearly one third of the causes of death being cancer. I believe interpretations make sense for the data output on risk factors because the assumptions under Method A were met.

Conclusion.

The overall theme of the research paper was to theorize and implement two types of methods for fitting a Cox Proportional Hazards Model to cause-specific survival data. The methods involved the assumption of two different failure types and any number of non-failure covariates. I appreciated the theory and challenge on determining the derivations behind the formulas presented in the paper. I also found the two applications to be very good at illustrating when to use one, both, or neither methods in accordance to following the assumptions. I found this topic to be very interesting and learned a lot!