THE COX PROPORTIONAL HAZARD MODEL

ROGER M. COOKE

1. Introduction

D.R. Cox's article in the Journal of the Royal Statistical Society "Regression Models and Life-Tables" (Series B (Methodological), Vol. 34, No. 2 (1972), pp. 187-220) introducing the proportional hazard model, is one of the most famous papers in statistics. According to Google Scholar as of February 2007 it has received over 8500 citations. The volume of literature on the Cox proportional hazard model is wide and deep ([18]). The theoretical foundations were secured years later with the development of the counting process approach (see [2, 3, 13]).

What explains the unparalleled success of the Cox model? This contribution attempts to give insight into the reasons for its success and also its limitations. The presentation is quasi formal, ideas are introduced and motivated from sampling considerations rather than from theorems.

The guiding idea is model adequacy. Suppose you want to explain human lifespan in terms of "covariates" that influence the lifespan. The first idea might be to build a regression model; you obtain data on how long people live, how much they smoke, their body mass index, their income, years of schooling, sex, race, exposure to pollution and whatever else you can think of and measure. You regress lifespan, or maybe log lifespan on these variables, in order to argue 'For someone in your situation, if you reduce smoking by A you can expect to live B years longer.' But first the question of model adequacy must be addressed, how much variance in the lifespan data do you explain with your model? If you are told someone's values on all these explanatory variables, how much better can you predict his/her lifespan than without this knowledge? If the answer is 'not very much' then your predictions shouldn't carry much persuasive force.

Or should they? Cox's key idea was to regress, not the lifespans themselves, but the *failure* or *hazard rates* onto explanatory variables. However, we don't observe failure rates, just failures. How do we minimize mean square difference between an observed and predicted failure rates? We don't; there is no 'residual in this sense. The beauty of Cox regression is that it provides another way of finding optimal estimates of the covariates' coefficients. Very roughly, we find values of the covariate coefficients such that the people with shorter lifespans tend to have larger failure rates. The method is called 'maximization of partial likelihood'. It works when we can factorize the failure rate, or its time integral, the cumulative hazard function, into a time independent part depending on covariates and a time dependent part independent of covariates. That explains the cohabitation of the terms "proportional hazard model" and "partial likelihood".

Key words and phrases. Competing risk, Relative risk, Cox proportional hazard, Censoring, Model adequacy, omitted covariates.

Is it all too good to be true? The goal of this contribution is to enable the reader to judge this for him/herself [1].

2. Proportional Hazard model

To simplify the presentation, we consider the case of time-invariant covariates X, Y, Z without censoring and without ties. We consider data to be generated by the following hazard rate

$$h(X,Y,Z) = \Lambda_0(t)e^{XA+YB+ZC} \tag{2.1}$$

where the cumulative baseline hazard function $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ is the time integral of the baseline hazard rate λ_0. The covariates (X, Y, Z) are considered as random variables. The coefficients (A, B, C) and the baseline hazard Λ_0 will be estimated from life data. If this hazard function holds, then for an individual with covariate values (x,y,z) the survivor function is

$$e^{-h(x,y,z)}. \tag{2.2}$$

Suppose we observe times of death $t_1, ...t_n$ such that $t_i < t_j$ for $i < j$. Let the covariates for the individual dying at time t_i be denoted (x_i, y_i, z_i). The coefficients A, B, C are estimated by maximizing the *partial likelihood*

$$\prod_{i=1}^{N} \frac{e^{x_i A+y_i B+z_i C}}{\sum_{j\geq i}^{n} e^{x_j A+y_j B+z_j C}} \tag{2.3}$$

Note that the times of death t_i do not appear in (2.3). The intuitive explanation is as follows. *Given* that the first death in the population occurs at time t_1, the probability that it happens to individual 1 is $\frac{e^{x_1 A+y_1 B+z_1 C}}{\sum_{j\geq 1}^{n} e^{x_j A+y_j B+z_j C}}$. After individual 1 is removed from the population, the same reasoning applies to the surviving population; *given* that the second time of death t_2, the probability that it happens to individual 2 is $\frac{e^{x_2 A+y_2 B+z_2 C}}{\sum_{j\geq 2}^{n} e^{x_j A+y_j B+z_j C}}$, and so on. Kalbfleisch and Prentice ([16]) note that for constant covariates, (2.3) is the likelihood for the *ordering* of times of death. The baseline hazard can be estimated from the data as described in ([16]p.114).

3. Model adequacy

With (x, y, z) fixed and T random, *and* with constant baseline hazard rate scaled to one, the survivor function (2.2), considered as a function of the random variable T is uniformly distributed on [0 , 1], that is

$$T \sim -ln(U)/h. \tag{3.1}$$

where U is uniform on $[0, 1]$. As this holds for each individual in the population $i = 1...N$. If we order the values

$$e^{-t_i e^{x_i A+y_i B+z_i C}}; i = 1...N \tag{3.2}$$

[1]The author gratefully acknowledges many helpful comments from Bo Lindqvist

and plot them against their number, the points should lie along the diagonal if the proportional hazard model is true with coefficients A, B, C and constant baseline hazard rate. (3.2) are the exponentials of the Cox Snell residuals; equal up to a constant to the Martingale residual, used in the counting process approach. The Cox Snell residuals are exponentially distributed if the model is correct.

This would provide an easy heuristic check of model adequacy if the baseline hazard were indeed known to be constant and scaled to one. However, if the baseline hazard is also estimated from the data, then this simple test does not apply.

Testing model adequacy for the Cox model is not straightforward: This is a sampling of statements found in the literature regarding model evaluation: "it is not apparent what kinds of departures one would expect to see in the residuals if the model is incorrect, or even to what extent agreement with the anticipated line should be expected" ([16], p128). "For most purposes, you can ignore the Cox-Snell and martingale residuals. While Cox-Snell residuals were useful for assessing the fit of the parametric models in Chapter 4, they are not very informative for the Cox models estimated by partial likelihood" ([1], p 173). "Unfortunately, this distribution theory [of the Cox Snell residuals as exponentially distributed] has not proven to be as useful for model evaluation as the theory derived from the counting process approach". ([14] p. 202), "there is not a single, simple, easy to calculate, useful, easy to interpret measure [of model performance] for a proportional hazards model." ([14] p. 229). "the martingale residuals can not play all the roles that linear model residuals do; in particular the overall distribution of the residuals does not aid in the global assessment of fit." ([20] p 81). In many important studies, model adequacy is not examined, and only individual coefficients for the covariate of interest are reported, with Wald confidence bounds (eg [12, 19]). The coefficients are used to compute relative risk, and form the basis of (dis)utility calculations for different risk mitigation measures.

It may well arise that data generated with a constant baseline hazard appear to acquire a time dependent baseline hazard as a result of missing covariates. Letting $\hat{\diamond}$ denote values estimated from the data, we may well find that the values

$$e^{-\hat{\Lambda}_0(t_i)e^{x_i\hat{A}+y_i\hat{B}+z_i\hat{C}}}; i = 1...N \tag{3.3}$$

plot as uniform, while the estimates do not resemble the values which generated the data. In particular, this may arise in the case of missing covariates. We identify some covariates but many others may not be represented in our model. For example, in considering the influence of airborne fine particulate matter on non-accidental mortality [12, 19], covariates like smoking, sex, age, socio-economic status, air quality, and weather are studied. However time to death is obviously influenced by myriad other factors like occupation, genetic disposition, stress, disease prevalence, medical care, diet, alcohol consumption, home environment (eg radon), travel patterns, etc.etc.

The following type of simple numerical experiment, which the reader may verify for him/herself will illustrate the problems with model adequacy[2].

(1) Choose coefficients (A, B, C), choose a constant baseline hazard rate scaled to one, and choose a distribution for (X, Y, Z);

[2]The following simulations were performed with EXCEL and checked with S+.

(2) Sample independently 100 values of (X, Y, Z) and 100 values from the uniform distribution on $[0, 1]$; compute failure times using (3.1);
(3) Estimate the coefficients by maximizing (2.3), and estimate the baseline hazard.

This procedure does not require that the distributions of the covariates be centered at their means; indeed, centering is not standard procedure in applications. However, the uniform distribution on [-1,1] used here is centered.

Let the model (2.1) be termed h_{XYZ}. To study the effects of model incompleteness estimate the coefficient A with a model h_{XY} using only covariates X and Y, and with a model h_X using only covariate X. For each of the models h_{XYZ}, h_{XY} and h_X, we repeat the above procedure 100 times with the same values for (A, B, C), with (X, Y, Z) sampled independently from the (centered) uniform distribution on $[-1, 1]$; however, we change (2.3) so as to estimate A in the models h_{XY} and h_X. Figure 1 plots the ordered estimates of coefficient A.

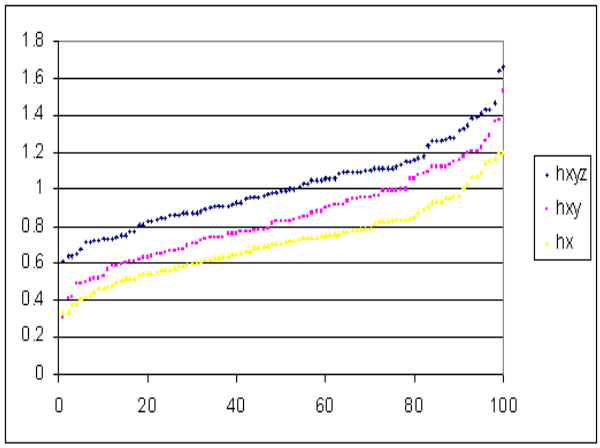

FIGURE 1. 100 ordered estimates of A for h_{XYZ}, h_{XY}, h_X, $X, Y, Z \sim U[-1, 1], (A, B, C) = (1, 1, 1)$; each estimate based on 100 samples

Evidently, the models h_{XY} and h_X tend to underestimate the coefficient A. A theoretical explanation of this underestimation is given in [5, 17]. Hougaard [15] shows that missing covariates sometimes result in non-proportional hazard functions which could be detected with sufficient replications. He concludes that "the regression coefficients cannot be evaluated without consideration of the variability of neglected covariates. This is a serious drawback of relative risk models"(p.92), leading him to promote accelerated failure or frailty models. The tendency to underestimate becomes more pronounced in Figure (2), where the missing covariate Z has coefficient $C = 5$. In spite of this, the ordered values of (3.3) plot along the diagonal, as shown in Figure (3). If we knew that the data was created with $\widehat{\lambda}_0(t) \equiv 1$, then we may impose this constraint on the survivor functions. From Figure (4) we see that uniformity is lost for the incomplete models hy_{XY}, h_X; but not for the complete model h_{XYZ}. This would provide an excellent diagnostic for

completeness if we had a priori knowledge of the baseline hazard rate; unfortunately in practice we do not have this knowledge. We can, however, find another diagnostic (see below).

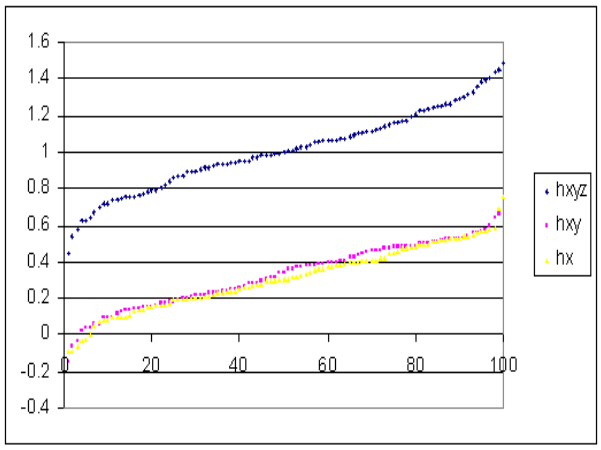

FIGURE 2. 100 ordered estimates of A for h_{XYZ}, h_{XY}, h_X, $X, Y, Z \sim U[-1, 1], (A, B, C) = (1, 1, 5)$ each estimate based on 100 samples

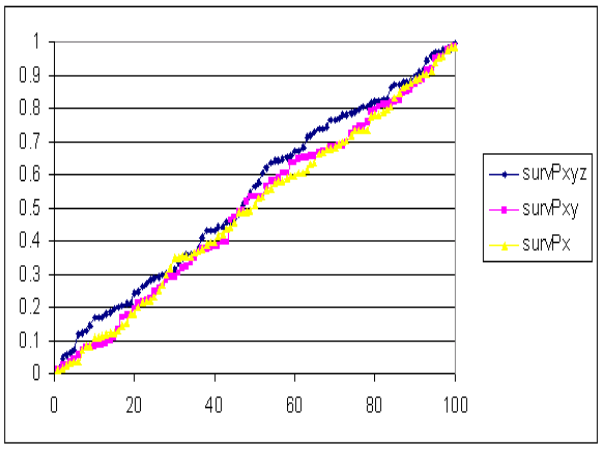

FIGURE 3. Ordered values of (3.3) for h_{XYZ}, h_{XY}, h_X, $X, Y, Z \sim U[-1, 1], (A, B, C) = (1, 1, 5)$

Figure (5) shows the Wald 95% confidence bounds for A in model h_X, in each of the 100 repetitions of the experiment whose estimates are shown in Figure (2). These bounds are derived assuming asymptotic normality of the Wald statistic

$$\frac{\hat{A} - A}{\sigma_A}$$

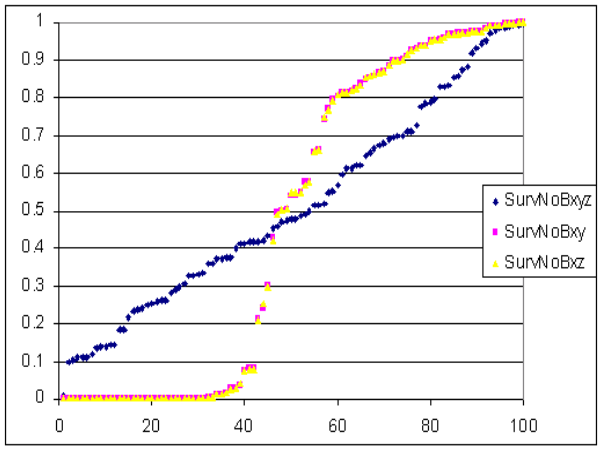

FIGURE 4. Ordered values of (3.3) for h_{XYZ}, h_{XY}, h_X, $X, Y, Z \sim U[-1,1], (A,B,C) = (1,1,5)$ with $\widehat{\Lambda}_0(t) \equiv t$;

where $\hat{A}$ is the estimate of A and σ_A is derived from the observed information matrix. If the likelihood function is correct, then the Wald statistic is asymptotically standard normal. In as much as these 95% confidence bands contain the true value $A = 1$ in only 7% of the cases, the wisdom of stating such confidence bounds when model adequacy cannot be demonstrated may be questioned.

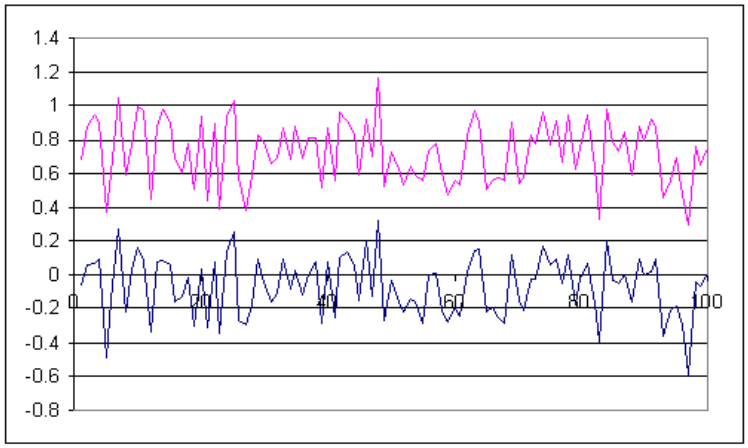

FIGURE 5. Wald 95% confidence bounds for A with model , h_X of Figure (2); each estimate based on 100 samples

The models h_{XY} and h_X are clearly incorrect and mis-estimate the covariate A. Relative risk coefficients based on these models would be biased. Without a priori knowledge of the baseline hazard function, their incorrectness cannot be diagnosed using Cox-Snell or Martingale residuals, echoing the statements cited at the beginning of this section. The problem is that the lack of fit caused by missing covariates is compensated in the estimated baseline hazard function.

This observation suggests that we might detect lack of fit in the covariates by comparing the estimated baseline hazard function with the population cumulative hazard function, that is, minus the natural logarithm of the population survivor function. From (3.3) it is evident that adding a constant to any covariate is equivalent to multiplying the baseline hazard by a constant. We therefore standardize the covariates by centering their distributions on the means (the distributions here already centered). This standardization can be motivated as follows. If the covariates X, Y and Z have no effect, then their means in the set R_i of individuals at risk at the time t_i of death of the $i-th$ individual, are just their means $\bar{x}, \bar{y}, \bar{z}$. We write

$$\sum_{j \in R_i} e^{x_j A + y_j B + z_j C} \sim \sum_{j \in R_i} (1 + x_j A)(1 + y_j B)(1 + z_j C).$$

If X,Y and Z are independent with mean zero then the mean of the right hand side is just $\#R_i$. The Breslow estimate [4] of the baseline hazard rate is

$$d\Lambda_0(t_i) = [\sum_{j \in R_i} e^{x_j A + y_j B + z_j C}]^{-1}$$

and the population cumulative hazard rate at t_i is $[\#R_i]^{-1}$. Hence, under these conditions, centering the covariates at their means, then replacing covariates with their mean value (zero), would equate the cumulative baseline hazard function and the population cumulative hazard function.

Figures (6, 7) show these comparisons for the two cases from Figures (1,2). Note the difference in survival times (horizontal axis); this is caused by the heavier loading of covariate Z in Figure (7). The Nelson Aalen estimator is used for the population cumulative hazard function.

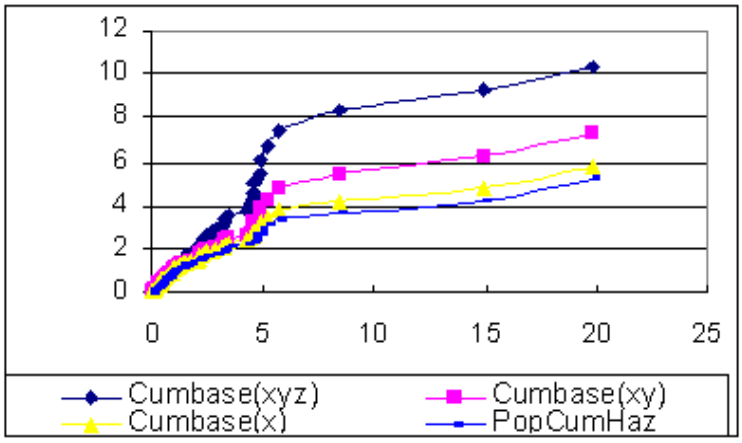

FIGURE 6. Cumulative population and baseline hazard functions for $h_{XYZ}, h_{XY}, h_X, X, Y, Z \sim U[-1, 1], (A, B, C) = (1, 1, 1)$

We see in Figure (7) that the cumulative baseline hazard functions for h_{XY} and h_X have moved closer to the population cumulative hazard, reflecting the heavier loading on the missing covariate Z.

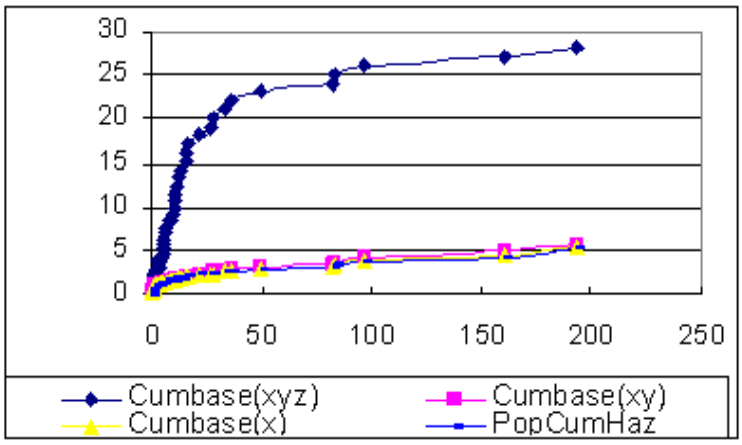

FIGURE 7. Cumulative population and baseline hazard functions for $h_{XYZ}, hy_{XY}, h_X, X, Y, Z \sim U[-1,1], (A,B,C) = (1,1,5)$

If a Cox model had *none* of the actual covariates, this would be equivalent to having zero coefficients on all covariates; and in this case the baseline hazard would coincide with the population cumulative hazard. A simple heuristic test of model adequacy would test the null hypothesis that the cumulative baseline hazard function is equal to the population cumulative hazard function. If the null hypothesis cannot be rejected, then using the Cox model would not be indicated. In Figures (8, 9) the asymptotic 2-sigma bands on the asymptotic variance of the Nelsen Aalen estimator of the population cumulative hazard function ([16], p.25) have been added to Figures (6, 7). We see that with this simple test we would fail to reject the null hypothesis for model h_X after 100 observations in both cases. The greater loading of missing covariate Z in Figure (9) causes the model h_{XY} to fail to reject the null hypothesis as well.

The more familiar partial likelihood ratio test calculates the test statistic G as twice the difference between the the log partial likelihood of the model containing the covariates and the log partial likelihood for the model not containing the covariates. G is asymptotically chi square distributed under the null hypothesis. The above test may have some advantage in that it does not appeal to partial likelihood. However, it is unable to detect the lack of fit in the model h_{XY} when $C = 1$.

We note that for all the results mentioned above, the covariates are independent. In practice independence is not usually checked, and not always plausible. Figure (10) shows 100 estimates of the coefficient A for the models h_{XYZ}, hy_{XY}, h_X where the covariates are uniformly distributed on $[0,1]$ with correlations $\rho(X,Z) = 0.98, \rho(Y,Z) = 0.41$. Whereas missing covariates produce under-estimation in the case of independence, we see that dependence in Figure(10) produces over-estimation. Note also that the spread of estimates for the complete model h_{XYZ} is very wide.

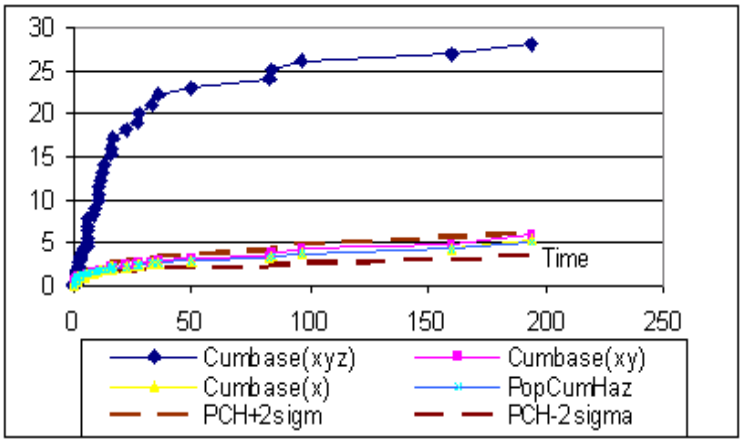

FIGURE 8. Cumulative population and baseline hazard functions for $h_{XYZ}, hy_{XY}, h_X, X, Y, Z \sim U[-1,1], (A,B,C) = (1,1,5)$ with 2-sigma confidence bands (dashed lines)

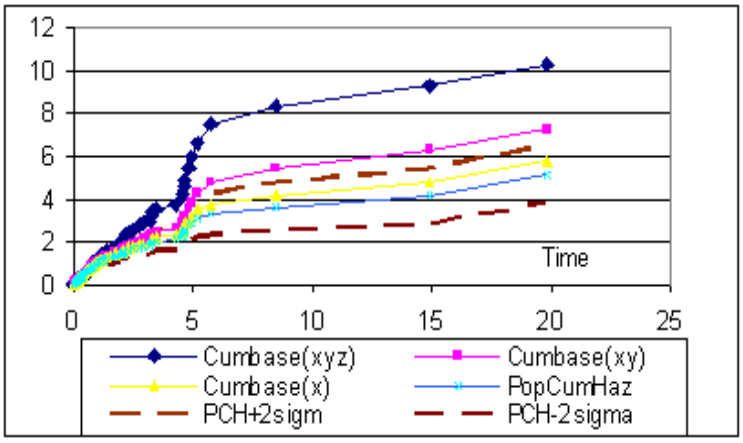

FIGURE 9. Cumulative population and baseline hazard functions for $h_{XYZ}, hy_{XY}, h_X, X, Y, Z \sim U[-1,1], (A,B,C) = (1,1,1)$ with 2-sigma confidence bands (dashed lines)

4. Censoring and Competing Risk

The discussion of model adequacy with the proportional hazard model is sometimes clouded by the role of censoring. The following statement is representative: "A perfectly adequate model may have what, at face value, seems like a terribly low R^2 due to a high percent of censored data" ([14] p. 229). The reference to R^2 must be taken as metaphorical. The proportional hazard model proposes a linear regression of the log hazard function. The hazard function is not observed, and hence a measure of the difference between observed and predicted values, like R^2

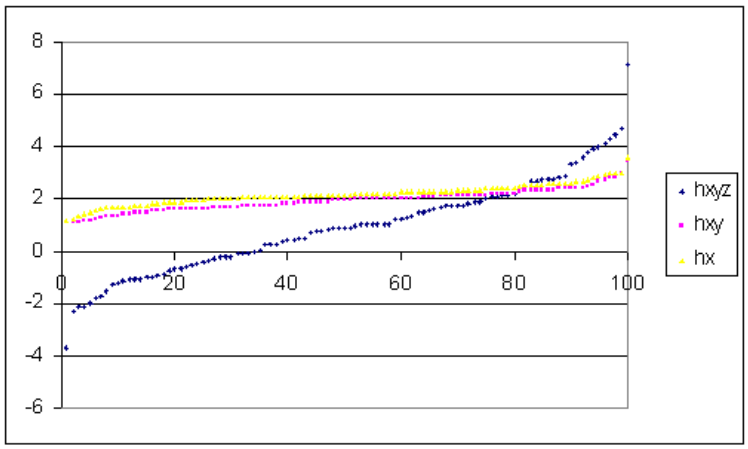

FIGURE 10. 100 ordered estimates of A for $h_{XYZ}, hy_{XY}, h_X, X, Y, Z \sim U[0,1], (A,B,C) = (1,1,1)$ with dependent covariates

is not meaningful. The point is that the ability of a proportional hazard model to "explain the data" might be obscured by censoring.

Right censoring of course is a form of competing risk. In this section we review some recent results from the theory of competing risk, and indicate how they may yield diagnostic tools in proportional hazard modelling. In the competing risks approach we model the data as a sequence of i.i.d. pairs (T_i, δ_i), $i = 1, 2, \ldots$. Each T is the minimum of two or more variables, corresponding to the competing risks. We will assume that there are two competing risks, described by two random variables D and C such that $T = \min(D, C)$. D will be time of death which is of primary interest, while C is a censoring time corresponding to termination of observation by other causes. In addition to the time T one observes the indicator variable $\delta = I(D < C)$ which describes the cause of the termination of observation. For simplicity we assume that $P(D = C) = 0$.

It is well known (Tsiatis[21]) that from observation of (T, δ) we can identify only the subsurvivor functions of D and C:

$$\begin{aligned} S_D^*(t) &= P(D > t, D < C) = P(T > t, \delta = 1) \\ S_C^*(t) &= P(C > t, C < D) = P(T > t, \delta = 0), \end{aligned}$$

but not in general the true survivor functions of D and C, $S_D(t)$ and $S_C(t)$. Note that $S_D^*(t)$ depends on C, though this fact is suppressed in the notation. Note also that $S_D^*(0) = P(D < C) = P(\delta = 1)$ and $S_C^*(0) = P(C < D) = P(\delta = 0)$, so that $S_D^*(0) + S_C^*(0) = 1$.

The conditional subsurvivor functions are defined as the survivor functions conditioned on the occurrence of the corresponding type of event. Assuming continuity of $S_D^*(t)$ and $S_C^*(t)$ at zero, these functions are given by

$$\begin{aligned} CS_D^*(t) &= P(D > t | D < C) = P(T > t | \delta = 1) = S_D^*(t)/S_D^*(0) \\ CS_C^*(t) &= P(C > t | C < D) = P(T > t | \delta = 0) = S_C^*(t)/S_C^*(0). \end{aligned}$$

Closely related to the notion of subsurvivor functions is the probability of censoring beyond time t,

$$\Phi(t) = P(C < D|T > t) = P(\delta = 0|T > t) = \frac{S_C^*(t)}{S_D^*(t) + S_C^*(t)}.$$

This function has some diagnostic value, aiding us to choose among competing risk models to fit the data. Note that $\Phi(0) = P(\delta = 0) = S_C^*(0)$.

As mentioned above, without any additional assumptions on the joint distribution of D and C, it is impossible to identify the marginal survivor functions $S_D(t)$ and $S_C(t)$. However, by making extra assumptions, one may restrict to a class of models in which the survivor functions are identifiable. A classical result on competing risks [21, 22] states that, assuming independence of D and C, we can determine uniquely the survivor functions of D and C from the joint distribution of (T, δ), where at most one of the survivor functions has an atom at infinity. In this case the survivor functions of D and C are said to be identifiable from the censored data (T, δ). Hence, an independent model is always consistent with data.

If the censoring is assumed to be independent then the survivor function for T, the minimum of D and C, can be written as

$$S_T(t) = S_D(t)S_C(t) \tag{4.1}$$

If we assume that D obeys a proportional hazard model, and that the censoring is independent, then we may estimate the coefficients by maximizing the partial likelihood function adapted to account for censoring:

$$\prod_{i \in D_N} \frac{e^{x_i A + y_i B + z_i C}}{\sum_{j \geq i}^{n} e^{x_j A + y_j B + z_j C}} \tag{4.2}$$

where D_N is the subset of observed times $t_1, ...t_N$ at which death is observed to occur, and j runs over all times corresponding to death or censoring.

If we now substitute the survivor function with estimated coefficients into (4.1), and use the familiar Kaplan Meier estimator for S_C, then we may apply the ideas of the previous section to assess model adequacy [8].

5. Conclusions

The user of the Cox proportional hazard model may be assured that *if* the model is correct, then the the coefficient estimates will converge to the true values, and the confidence bands will accurately reflect sampling fluctuations around these true values. If the model is *not* correct, however, then the coefficients' estimates may not be correct, and the confidence bands may give a misleading picture. If the covariates are independent, and if one or more covariates is omitted, then the coefficients of the other covariates will be underestimated in absolute value. If there is dependence between missing and included covariates, then nothing can be said about the direction of the error.

How can we check whether the model is correct? This remains the difficult point. We can test whether covariates are significantly different from zero. However, when covariates pass this test, it does not mean that they will be estimated correctly.

References

[1] Allison, P.D. (2003) "Survival Analysis Using SAS a practical guide" SAS Publishing, Cary.

[2] Andersen, P.K. and Gill, R.D. (1992) "Cox's regression model for counting processes: a large sample study" *The Annals of Statistics* vol. 10, no. 4, 1100-1120.

[3] Andersen, P.K., Borgan, O. Gill, R.D. and Keiding, N. (1992) *Statistical Models Based on Countinhg Processes* Springer-Verlag, New York,

[4] Breslow, N. (1972) Discusson on Professor cox's paper, JRSS, B , 34, 216-217.

[5] Bretagnolle, J. and Huber-Carol, C. (1988)Effects of omitting covariates in Cox's Model for susrvival data"Scand. J. Statist. 15, 125-138.

[6] C. Bunea, R.M. Cooke, and B. Lindqvist "Competing risk perspective over reliability data bases" (2003), Mathematical and Statistical Methods in Relibiity (B.H. Lindqvist and K.A.Doksum eds), World Scientific Publishing, Singapore, p 355-370.

[7] C. Bunea, R.M. Cooke and B. Lindqvist,(2002) Maintenance study for components under competing risks, in *Safety and Reliability, 1st volume*, (European Safety and Reliability Conference, ESREL 2002, Lyon, France, 18-21 March 2002), pp. 212-217.

[8] Cooke R.M. and Morales, O.(2006) "Competing Risk and the Cox proportional hazard model", Journal of Statistical Planning and Inference vol. 136, no 5 , pp. 1621-1637, 2006.

[9] R.M. Cooke, The design of reliability databases Part I and II,(1996) *Reliability Engineering and System Safety*, **51**, 137-146 and 209-223.

[10] R.M. Cooke and T.J. Bedford, Reliability databases in perspective (2002), *IEEE Transactions on Reliability*, vol. 51, no 3, (September 2002), pp. 294-310.

[11] D.R. Cox, (1972). "Regression Models and Life-Tables" Royal Statistical Society, Series B, Vol. 34, No. 2, 1972, pp. 187-220.

[12] Dockery, D., III, C. A. P., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. (1993). "An association between air pollution and mortality in six U.S. cities." New England Journal of Medicine, 329, 1753-1759.

[13] Fleming, T.R. and Harrington D.P. (1991) Counting Processes and Survival Analysis Wilely, New York.

[14] Hosmer, D.W. and Lemeshow, S. (1999) Applied Survival Analysis, Wiley, New York

[15] Hougaard, P. (2000) *Analysis of Multivariate Survival Data* SApringer-Verlag, New York.

[16] Kalbfleisch, J.D. and Prentice, R.L. (2002) The Statistical Analysis of Failure Time Data, second edition. Wiley New York.

[17] Keiding, N., Andersen, P.K., Klein, J.P. (1997) "The role of frailty time models in describing heterogeneity due to omitted covariates"*Statistic's in Medicine*, vol. 16, 215-224.

[18] Oakes, D. (2001) "Biometrika Centenary: Survival analysis" Biometrika 88 99-142.

[19] Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D., Evans, J. S., Speizer, F. E., and Heath, C. W. (1995). "Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults." American Journal of Respiratory and Critical Care Medicine, 151, 669-674.

[20] Therneau, T.M. and Grambsch, P.M. (2000) Modeling Survival Data, Springer, New York.

[21] A. Tsiatis, A nonidentifiability aspect of the problem of competing risks, (1975) *Proceedings of the National Academy of Sciences, USA*, **72**, 20-22.

[22] van der Weide, J.A.M. van der, and Bedford T., "Competing risks and eternal life" (1998), in *Safety and Reliability (Proceedings of ESREL'98)*, S. Lydersen, G.K. Hansen, H.A. Sandtorv (eds), Vol 2, 1359-1364, Balkema, Rotterdam.

Resources for the Future, Washington DC, and Delft University of Technology, Department of Mathematics, Delft, The Netherlands

E-mail address: cooke@rff.org