Diagnostics for Cox's Proportional Hazards Model

C. Caroni

National Technical University of Athens, Athens, Greece

Abstract: Although Cox's proportional hazards model is applied so often, diagnostic methods are not well known to non-expert users. Latest results suggest that individually scaled Schoenfeld residuals, dfbetas and added variable plots should be provided for assessing the model. Most of these methods are absent from widely used packages.

Keywords and phrases: Lifetime data, Cox's proportional hazards model, diagnostic methods, residuals, influence, added variable plots

3.1 Introduction

A large number of diagnostic methods for model checking have been proposed over the years for Cox's proportional hazards (PH) regression model, but they are still not used as widely as they should be. The purpose of this paper is to present an up-to-date review of the model checking procedures that seem to have the greatest practical importance and deserve to receive the most attention. The PH model [Cox (1972)] specifies

$$\lambda(t; x) = \lambda_0(t)e^{\beta'x}$$

for the hazard function for the failure time $T$ associated with a $p$-dimensional vector of covariates $x$, given an unknown underlying hazard function $\lambda_0$ and a vector of coefficients $\beta$ which must be estimated. Articles dealing with this model often start with the observation that in many fields it has become the standard way of analyzing lifetime data with covariates. This makes Cox's model one of the big success stories of statistics. At the same time, once a model has reached the status of being the standard, there is a great danger that people will apply it because everyone else does, and that referees of scientific papers will accept it for the same reason. Yet it is a statistical model, therefore
it contains assumptions that might not be suitable for the problem in question. The fact that it is a semi-parametric model does not change this. It is therefore important that diagnostic methods become widely known.

Previous work along the same lines includes a review by Hess (1995) of graphical methods for examining the PH assumption and an empirical comparison of statistical tests for PH by Ng’andu (1997). The more modern books, such as Klein and Moeschberger (1997), Therneau and Grambsch (2000) and Harrell (2001) also contain much relevant material, although heavily linked in some cases to particular packages such as S-Plus.

From the hazard function of the PH model, we obtain the survivor function

\[ S(t; x) = \exp\{-\Lambda_0(t) e^{\beta' x}\} \]

where \( \Lambda_0 \) is the cumulative hazard corresponding to \( \lambda_0 \). Hence

\[ \ln\{-\ln S(t; x)\} = \ln \Lambda_0 + \beta' x. \]

This means that any survivor function \( S(t; x) \), when plotted on the complementary log-log scale, differs from \( \ln \Lambda_0 \) by the constant amount \( \beta' x \) throughout its duration. Hence any two such functions, \( S(t; x_1) \) and \( S(t; x_2) \) for different values of the covariate vector \( x \), will be parallel. This gives the basic means of inspecting the PH assumption, maybe the only one suggested in the more elementary books such as Parmar and Machin (1995):

- obtain Kaplan-Meier estimates \( \hat{S}(t; x) \) for various selected \( x \)
- examine visually whether the curves \( \ln\{-\ln \hat{S}(t; x)\} \) against \( t \) (or a function of \( t \)) are parallel for different \( x \).

The point of using a function of \( t \) is to make it easier to assess the graph. If \( \ln t \) is used, then plots of \( \ln\{-\ln \hat{S}(t; x)\} \) will be straight lines when lifetimes follow the Weibull distribution. It is probably easier for the eye to assess straight lines than curves.

The main drawback of this procedure is obvious. We need to have reliable estimates \( \hat{S} \) and this means having a large amount of data available for each chosen \( x \). This will only happen if there are few covariates and few distinct values (although continuous variates can be grouped), for example if there is just one covariate with a small number of levels. Therefore, we will not often be able to use this plot although, when we can, it is essential to look at it before doing the work of fitting the PH model.

We now move on to the more important diagnostic methods based on residuals and other information obtained from the fit of the PH model to the data.