Biostatistics (MATH11230)

Checking the proportional hazards assumption

Vanda Inácio

University of Edinburgh

Semester 1, 2022/2023

Checking the proportional hazards assumption

- Over the last decades, Cox proportional hazards (PH) model became, possibly, the most popular regression model for time to event/survival data and is widely used in medicine and beyond.
- But how to test the most famous assumption of Cox's model: proportional hazards?
- Remember that the assumption that hazards are proportional for a given risk factor say, gender, means that when we plot the (cumulative) hazard for females and the (cumulative) hazard for males, the two hazard curves are parallel.
- This means that for any point in time we can multiply the hazard for females and obtain the hazard for males.
- If this relation holds, the hazard ratio for females compared to males is a good one number summary of the relation.

Checking the proportional hazards assumption

- ↪ However, the proportional hazards assumption is not always realistic and in particular, in medicine, it has been challenged over the last years by the emergence of new types of treatments having different mechanisms of action.

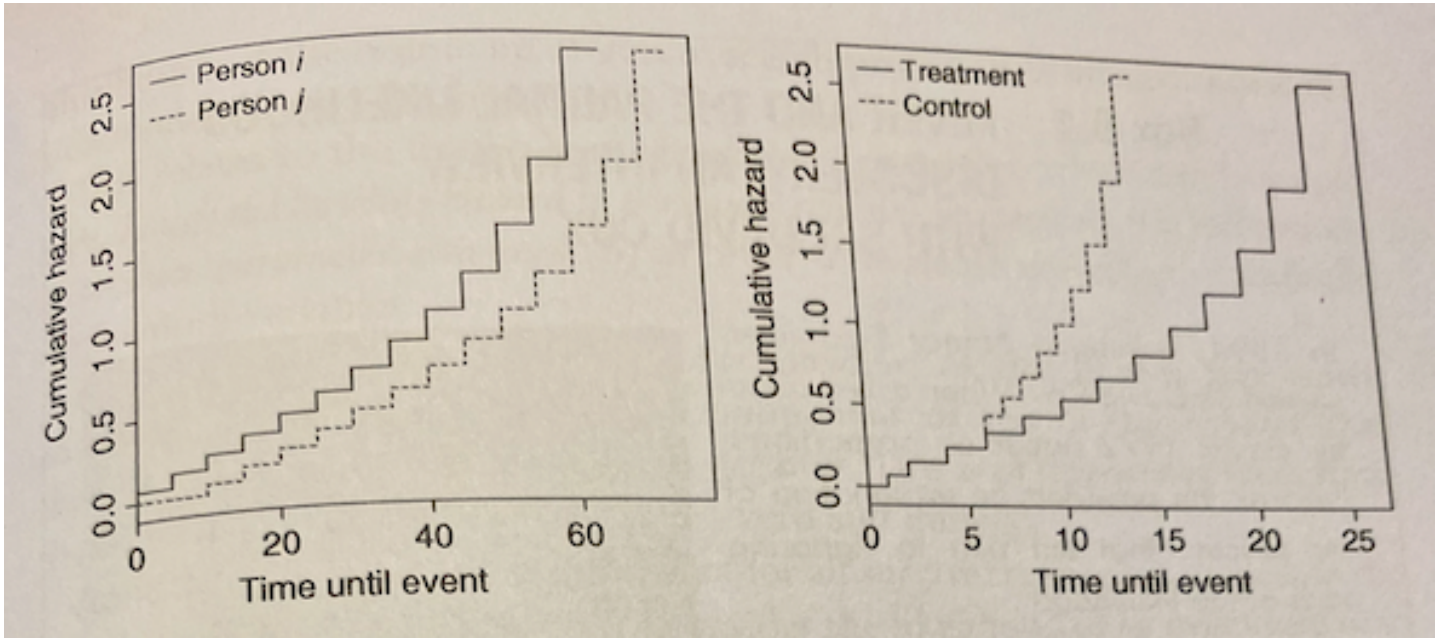
- ↪ Consider, for example, the introduction of a new type of treatment in a clinical trial. Two groups of subjects are randomly divided intro the standard and new treatment groups.

- ↪ It may be that both groups have similar survival rates in the first few weeks, since they were randomised and all should theoretically have the same risk.

- ↪ Through the progression of time, however, the impact of the new treatment could mean a progression to death at a slower rate in this group than that of those in the standard treatment group. In this case, the hazards would in this case not be parallel for all time.

Checking the proportional hazards assumption

↪ The below figure from Mills (2011, p. 89) illustrates the case of proportional hazards (left) and nonproportional hazards (right).



Checking the proportional hazards assumption

Comparing predicted survival curves based on Cox model and based on the Kaplan-Meier estimator

- We have already learned how to compute predicted survival curves from the Cox model.
- It is then possible to assess violations of the PH assumption by comparing survival curve estimates based on the Cox PH model with estimates computed independently of the model.
- For a single binary risk factor, we can compare the Kaplan–Meier estimate for each group with the estimates computed from the Cox model.
- Clear departures of the two estimates would provide evidence against the PH assumption.
- This idea extends to multiple binary or categorical covariates. Continuous risk factors can also be considered but must first be categorised. It is usually recommended to consider a reasonably low number of categories, keeping sufficient observations in each category.
- Note that the resulting estimated curves will be step functions, due to the nature of the underlying estimators of the survival/cumulative hazard functions, and might be rather poorly estimated for the levels of the risk factor with less observations.

Checking the proportional hazards assumption

Comparing predicted survival curves based on Cox model and based on the Kaplan-Meier estimator

- In this approach, it is not always possible to judge whether discrepancies between model based (from Cox PH model) and non-model based (Kaplan-Meier) survival estimates result from sampling fluctuations or from real trends.

- The presentation of confidence intervals may help.

- Calculation of the confidence intervals for the Kaplan–Meier estimates is readily available in R but confidence intervals for the survival curve estimates derived from the Cox model are not directly available in R and may involve fairly complex calculations. Of course, we can always resort to the bootstrap to obtain confidence bands.

- Even if calculation of the intervals were trivial, plotting the survival curve estimates and associated confidence intervals would lead to a busy plot. A remedy would be to plot the confidence intervals on a separate plot (possibly a plot for the lower bands and another one for the upper bands).

Checking the proportional hazards assumption

The log cumulative hazard plot

- We have also already learned that under the proportional hazards model we have that $S(t \mid \mathbf{x}) = S_0(t)^{\exp\{\mathbf{x}'\beta\}}$ and thus the survival curves are powers of one another.
- Therefore, a rudimentary check of the PH assumption is through inspection of the Kaplan-Meier curve estimates, since the crossing of survival curves indicates departure from the PH assumption.
- The PH assumption also implies that $H(t \mid \mathbf{x}_i) = H_0(t)\exp(\mathbf{x}_i'\beta)$.
- To see why, remember that the PH model is given by

$$h(t \mid \mathbf{x}_i) = h_0(t)\exp(\mathbf{x}_i'\beta).$$

- Integrating both sides from 0 up to t, yields

$$\begin{aligned}\int_0^t h(u \mid \mathbf{x}_i)\mathrm{d}u &= \int_0^t h_0(u)\exp(\mathbf{x}_i'\beta)\mathrm{d}u \\ \Rightarrow H(t \mid \mathbf{x}_i) &= H_0(t)\exp(\mathbf{x}_i'\beta),\end{aligned}$$

where $H(t \mid \mathbf{x}_i)$ and $H_0(t)$ are the cumulative hazard functions.

Checking the proportional hazards assumption

The log cumulative hazard plot

- Taking logarithms of each side of this equation, we get

$$\log H(t \mid \mathbf{x}_i) = \log H_0(t) + \mathbf{x}_i'\beta,$$

or, equivalently,

$$\log\{-\log S(t \mid \mathbf{x}_i)\} = \log\{-\log S_0(t)\} + \mathbf{x}_i'\beta.$$

- It thus follows that differences in the log cumulative hazard functions do not depend on time.

- This means that if the log cumulative hazard functions for individuals with different values of their explanatory variables are plotted against time, the curves so formed will be parallel if the proportional hazards model is valid.

- This approach only requires estimating $H(t \mid \mathbf{x}_i)$ or $S(t \mid \mathbf{x}_i)$ nonparametrically, for example, via the Nelson–Aalen or the Kaplan–Meier estimator, for the groups of individuals defined by the different values of the risk factor(s) under study.

Checking the proportional hazards assumption

The log cumulative hazard plot

- As for the previous approach, this approach is most useful when we have a limited number of risk factors and a sufficient number of observations per each level of the risk factors. Note that as before, when stratifying by several risk factors, the number of observations in each combination may be so low that it prevents accurate estimation of the corresponding survival/cumulative hazard function.

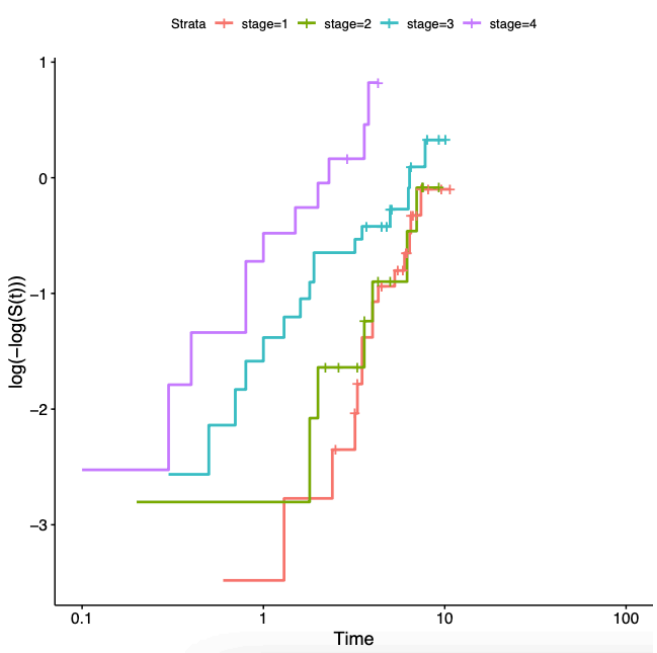
- To decide whether these curves are indeed parallel will obviously always be subject to some subjectivity.

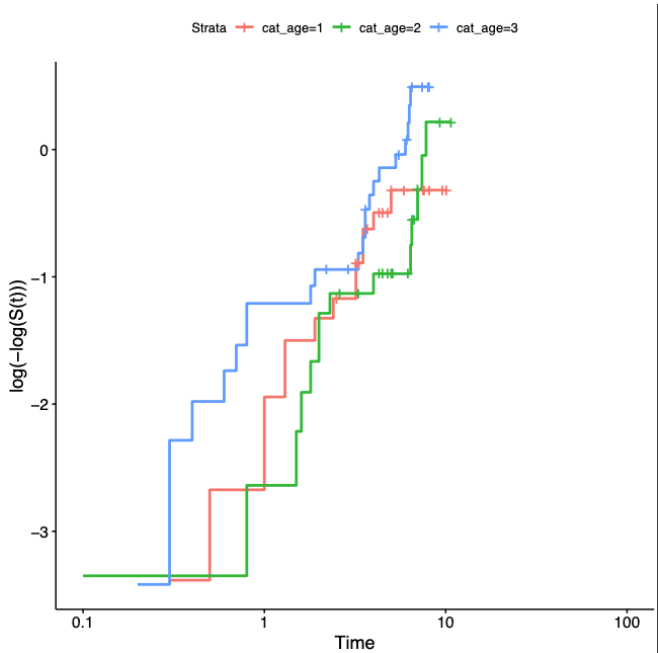
- A commonly adopted attitude is to assume that the proportional hazards assumption is satisfied unless these curves show a strong deviation from parallel curves.

Checking the proportional hazards assumption

The log cumulative hazard plot

↪ Investigating if the proportional hazards assumption is valid for the variables stage and age in the larynx cancer example (age has been categorised according to the 0.33 and 0.66 quantiles).





Checking the proportional hazards assumption

Partitioning the time axis and fitting Cox models to each interval

- ↪ Under this approach, follow-up time is partitioned into intervals, and separate PH models are fitted to each time interval. Of course, we assume that the PH assumption holds within each interval.
- ↪ One can then compare the coefficients (i.e., the log hazard rations) across intervals.
- ↪ For each interval, we exclude all subjects whose event/censoring time occurred before the start of the interval and we censor the event times for subjects who do not experience the event through the interval at the end of the interval.
- ↪ It is recommended that intervals be constructed so that there is roughly an equal number of events in each. The number of intervals should allow at least 10 or 20 events per interval.
- ↪ By plotting the log hazard ratio and its confidence limits versus the interval, one can assess the importance of a predictor as a function of follow-up time and learn how to model non PH using more complicated models containing predictor by time interactions

Checking the proportional hazards assumption

Schoenfeld residuals

- Another graphical option is to use the Schoenfeld residuals, also known as partial residuals.

- For the kth risk factor, $k = 1, \ldots, p$, and the ith individual, the Schoenfeld residual represents the difference between the observed value of x_{ik} and its conditional expectation given the risk set $R(y_i)$

$$r_{ik} = \delta_i(x_{ik} - \hat{a}_{ik}), \quad \hat{a}_{ik} = \frac{\sum_{l \in R(y_i)} x_{lk} \exp(\mathbf{x}_l' \widehat{\boldsymbol{\beta}})}{\sum_{l \in R(y_i)} \exp(\mathbf{x}_l' \widehat{\boldsymbol{\beta}})}.$$

- Note that Schoenfeld residuals will produce for each observation one value per risk factor included in the model.

- Non-zero values of these residuals only arise for uncensored observations.

- Moreover, if the largest observation in a sample of survival times is uncensored, the value of $\hat{a}_{ik}$ for that observation will be equal to x_{ik} and so $r_{ik} = 0$.

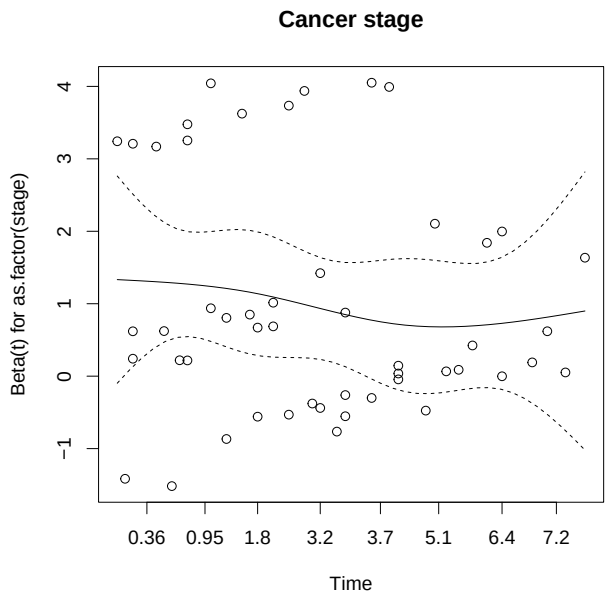
- These residuals sum to zero and, in large samples, they are uncorrelated to each other and have expected value zero.

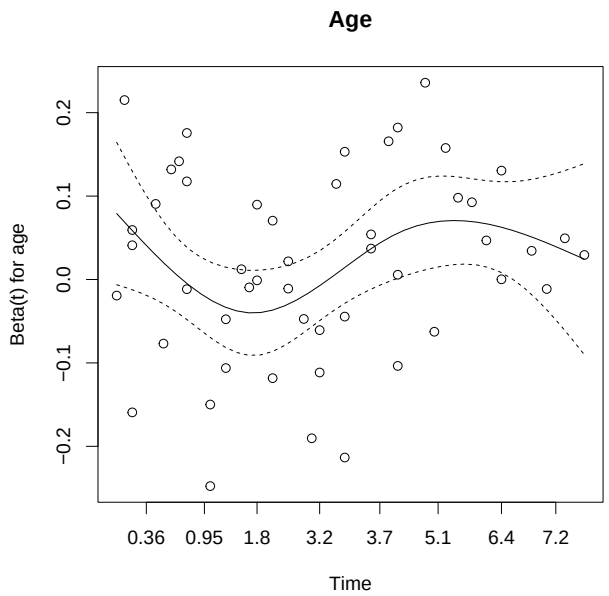
Checking the proportional hazards assumption

Schoenfeld residuals

- It has been shown that if the proportional hazards assumption is valid, Schoenfeld residuals should be flat and centred about zero.

- It has also been demonstrated that these residuals have to be independent of the time and so if we represent them ranked by its event time, this plot must not show any pattern.





Checking the proportional hazards assumption

Schoenfeld residuals

- ↪ Trends in the scatterplots of the Schoenfeld residuals are often difficult to ascertain, especiallywith binary covariates where there are only two horizontal bands of residuals present.
- ↪ Superposition of the results of a smoothing procedure improves the interpretability of the residual plots.
- ↪ Smoothing helps describe the pattern of dependence of the mean of a response variable y (in this case the residuals) as a function of a variable x (in this case time).
- ↪ In the previous plot the solid black line is a LOWESS (locally-weighted scatterplot smoothing) fit. The only assumption of this method is is that the mean of y at each value of x, $\mathbb{E}(y \mid x)$, varies as a smooth function of x.

Checking the proportional hazards assumption

Goodness-of-fit tests

- A second approach for assessing the PH assumption involves goodness-of-fit tests.
- To this end, different test have been proposed in the literature (Grambsch and Therneau 1994).
- We focus on the test proposed by Harrell (1986), a variation of a test originally proposed by Schoenfeld (1982). This is a test of correlation between the Schoenfeld residuals and event time.
- A correlation of zero indicates that the model meet the proportional hazards assumption (the null hypothesis).

Checking the proportional hazards assumption

Non-Proportional Hazards... and now what?

- ↪ A 'modest' violation of the proportional hazards assumption may not make a big difference for large datasets.
- ↪ What if the nonproportionality is large and real?
- ↪ Risk factors with nonproportional effects may be incorporated into the model as stratification factors rather than explanatory variables.
- ↪ The idea is to consider a categorical variable defining the 'strata' in the population and to consider a different (unspecified) baseline hazard function in each strata.
- ↪ So, the hazard for individual i in strata j, $i = 1, \ldots, n_j$ and $j = 1, \ldots, J$, with risk factor values x_{ij}, is now given by
$$h_{ij}(t) = h_{0j}(t) \exp(\mathbf{x}'_{ij}\beta).$$
- ↪ The strata can, for example, be defined based on the gender.
- ↪ While this may be convenient for some applications to overcome the nonproportional hazards issue, we loose all the information about the stratification factor in the sense that this model does not provide information on the effect of the stratification variable.