Chapter 8

Assessing the PH Assumption

So far, we've been considering the following Cox PH model:

$$\lambda(t, \mathbf{Z}) = \lambda_0(t)\ \exp(\beta\mathbf{Z}) = \lambda_0(t)\ \exp\left(\sum \beta_j Z_j\right)$$

where β_j is the parameter for the the j-th covariate (Z_j).

8.1 Important features of this model

(1) the baseline hazard depends on t, but not on the covariates $Z_1, ..., Z_p$

(2) the hazard ratio, i.e., $\exp(\beta\mathbf{Z})$, depends on the covariates $\mathbf{Z} = (Z_1, ..., Z_p)$, but not on time t.

Assumption (2) is what led us to call this a proportional hazards model. That's because we could take the ratio of the hazards for two individuals with covariates $\mathbf{Z}_i$ and $\mathbf{Z}_{i'}$, and write it as a constant in terms of the covariates.

8.2 Hazard Ratio

$$\begin{aligned}
\frac{\lambda(t,\mathbf{Z}_i)}{\lambda(t,\mathbf{Z}_{i'})} &= \frac{\lambda_0(t)\exp(\beta\mathbf{Z}_i)}{\lambda_0(t)\exp(\beta\mathbf{Z}_{i'})} \\
&= \frac{\exp(\beta\mathbf{Z}_i)}{\exp(\beta\mathbf{Z}_{i'})} \\
&= \exp[\beta(\mathbf{Z}_i - \mathbf{Z}_{i'})] \\
&= \exp[\sum \beta_j (Z_{ij} - Z_{i'j})] \;=\; \theta
\end{aligned}$$

In the last formula, Z_{ij} is the value of the j-th covariate for the i-th individual. For example, Z_{42} might be the value of GENDER (0 or 1) for the the 4-th person.

We can also write the hazard for the i-th person as a constant times the hazard for the i'-th person:

$$\lambda(t,\mathbf{Z}_i) \;=\; \theta\;\lambda(t,\mathbf{Z}_{i'})$$

Thus, the HR between two types of individuals is constant (i.e., $=\theta$) over time. These are mathematical ways of stating the proportional hazards assumption.

There are several options for checking the assumption of proportional hazards:

I. **Graphical**

   (a) Plots of survival estimates for two subgroups
   (b) Plots of $\log[-\log(\hat{S})]$ vs $\log(t)$ for two subgroups
   (c) Plots of weighted Schoenfeld residuals vs time
   (d) Plots of observed survival probabilities versus expected under PH model (see Kleinbaum, ch.4)

II. **Use of goodness of fit tests** - we can construct a goodness-of-fit test based on comparing the observed survival probability (from `sts list`) with the expected (from `stcox`) under the assumption of proportional hazards - see Kleinbaum ch.4

III. **Including interaction terms between a covariate and** t (time-dependent covariates)

How do we interpret the above?

Kleinbaum (and other texts) suggest a strategy of assuming that PH holds unless there is very strong evidence to counter this assumption:

- estimated survival curves are fairly separated, then cross
- estimated log cumulative hazard curves cross, or look very unparallel over time
- weighted Schoenfeld residuals clearly increase or decrease over time (you could fit a OLS regression line and see if the slope is significant)
- test for time $\times$ covariate interaction term is significant (this relates to time-dependent covariates)

If PH doesn't exactly hold for a particular covariate but we fit the PH model anyway, then what we are getting is sort of an average HR, averaged over the event times.

In most cases, this is not such a bad estimate. Allison claims that too much emphasis is put on testing the PH assumption, and not enough to other important aspects of the model.

8.3 Implications of proportional hazards

Consider a PH model with a single covariate, Z:

$$\lambda(t; Z) = \lambda_0(t)e^{\beta Z}$$

What does this imply for the relation between the survivorship functions at various values of Z?

Under PH,

$$\log[-\log[S(t; Z)]] = \log[-\log[S_0(t)]] + \beta Z$$

In general, we have the following relationship:

$$\begin{aligned}
\Lambda_i(t) &= \int_0^t \lambda_i(u)du \\
&= \int_0^t \lambda_0(u)\exp(\beta\mathbf{Z}_i)du \\
&= \exp(\beta\mathbf{Z}_i)\int_0^t \lambda_0(u)du \\
&= \exp(\beta\mathbf{Z}_i)\,\Lambda_0(t)
\end{aligned}$$

This means that the ratio of the cumulative hazards is the same as the ratio of hazard rates:

$$\frac{\Lambda_i(t)}{\Lambda_0(t)} = \exp(\beta\mathbf{Z}_i) = \exp(\beta_1 Z_{1i} + \cdots + \beta_p Z_{pi})$$

Using the above relationship, we can show that:

$$\begin{aligned}
\beta\mathbf{Z}_i &= \log\left(\frac{\Lambda_i(t)}{\Lambda_0(t)}\right) \\
&= \log\Lambda_i(t) - \log\Lambda_0(t) \\
&= \log[-\log S_i(t)] - \log[-\log S_0(t)] \\
\text{so}\quad \log[-\log S_i(t)] &= \log[-\log S_0(t)] + \beta\mathbf{Z}_i
\end{aligned}$$

Thus, to assess if the hazards are actually proportional to each other over time

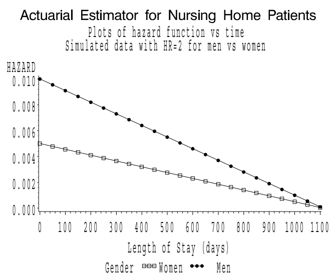
- calculate Kaplan Meier Curves for various levels of Z
- compute $\log[-\log(\hat{S}(t;Z))]$ (i.e., log cumulative hazard)
- plot vs log-time to see if they are parallel (lines or curves)

Note: If Z is continuous, break into categories.

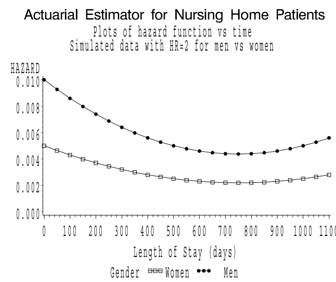
Question: Why not just compare the underlying hazard rates to see if they are proportional?

Here's two simulated examples with hazards which are truly proportional between the two groups:

Weibull-type hazard



U-shaped hazard

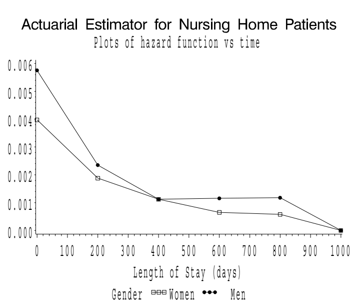


Reason 1: It's hard to eyeball these figures and see that the hazard rates are proportional - it would be easier to look for a constant shift between lines.

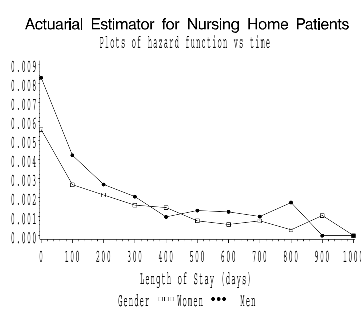
Reason 2: Estimated hazard rates tend to be more unstable than the cumulative hazard rate

Consider the nursing home example (where we think PH is reasonable). If we group the data into intervals and calculate the hazard rate using actuarial method, we get these plots:

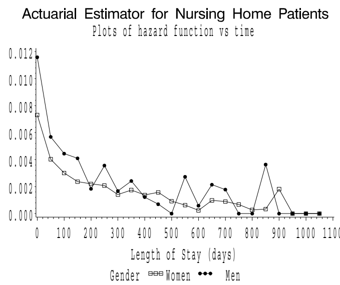
200 day intervals



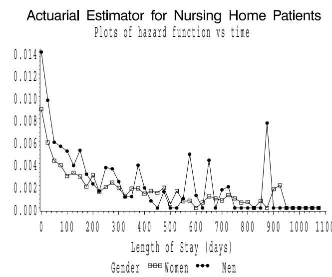
100 day intervals



50 day intervals



25 day intervals

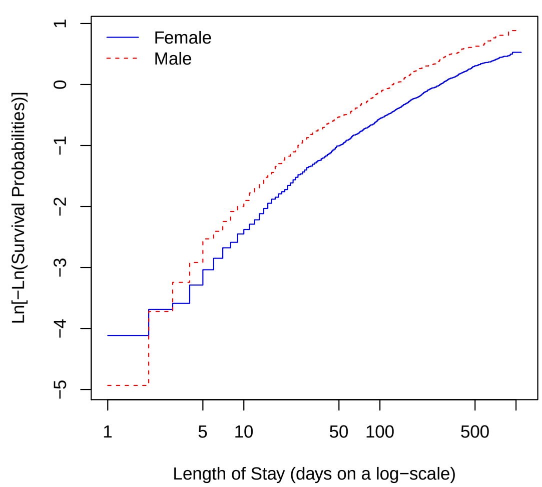


In contrast, the log cumulative hazard plots are easier to interpret and tend to give more stable estimates

8.3.1 Assessing proportionality with one covariate

Example: Nursing Home - gender

Figure 8.1: Assessing the PH assumption for the gender effect



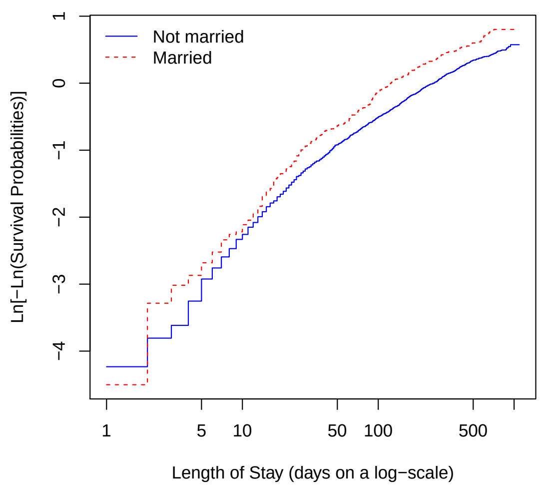

```
fitKMgen = survfit( Surv(los,fail) ~ gender,data = nurshome)

plot(fitKMgen, mark.time = F, fun = "cloglog",
     xlab = "Length of Stay (days on a log-scale)",
     ylab = "Ln[-Ln(Survival Probabilities)]",
```

```
     lty = 1:2,col = c("blue","red"))
legend("topleft",lty = 1:2,col = c("blue","red"),bty="n",
       legend = c("Female","Male"))
```

Example: Nursing Home - marital status

A similar situation is the case with respect to the effect of marital status; This is equivalent to comparing plots of the log cumulative hazard, $\log(\hat{\Lambda}(t))$,

Figure 8.2: Assessing the PH assumption for the effect of marital status



between the covariate levels, since

$$\Lambda(t) = \int_0^t \lambda(u; Z)du = -\log[S(t)]$$

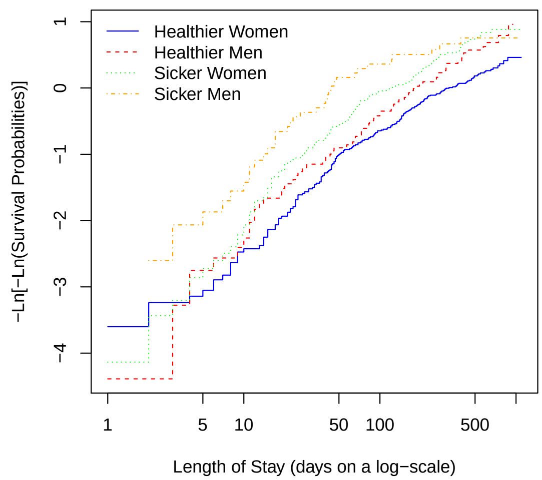
8.3.2 Assessing proportionality with several covariates

If there is enough data and you only have a couple of covariates, create a new covariate that takes a different value for every combination of covariate values.

Example: Health status and gender for nursing home

```
nurshome$hlthsex = c( 1*(nurshome$gender==0 & nurshome$health==2) +
                      2*(nurshome$gender==1 & nurshome$health==2) +
                      3*(nurshome$gender==0 & nurshome$health==5) +
                      4*(nurshome$gender==1 & nurshome$health==5) )
```

Figure 8.3: Log[-log(survival)] Plots for Health status*gender



If there are too many covariates (or not enough data) for this, then there is a way to test proportionality for each variable, one at a time, using the stratification option.

What if proportional hazards fails?

- do a stratified analysis
- include a time-varying covariate to allow changing hazard ratios over time
- include interactions with time

The second two options relate to time-dependent covariates, which is getting beyond the scope of this course.

We will focus on the first alternative, and then the second two options will be briefly described.

8.4 Stratified Analyses

Suppose:

- we are happy with the proportionality assumption on Z_1

- proportionality simply does not hold between various levels of a second variable Z_2.

If Z_2 is discrete (with a levels) and there is enough data, fit the following **stratified model**:

$$\lambda(t; Z_1, Z_2) = \lambda_{Z_2}(t) e^{\beta Z_1}$$

For example, a new treatment might lead to a 50% decrease in hazard of death versus the standard treatment, but the hazard for standard treatment might be different for each hospital.

A stratified model can be useful both for primary analysis and for checking the PH assumption.

8.4.1 Assessing PH Assumption for Several Covariates

Suppose we have several covariates ($\mathbf{Z} = Z_1, Z_2, \ldots\ Z_p$), and we want to know if the following PH model holds:

$$\lambda(t; \mathbf{Z}) = \lambda_0(t)\ e^{\beta_1 Z_1 + \ldots + \beta_p Z_p}$$

To start, we fit a model which stratifies by Z_k:

$$\lambda(t; \mathbf{Z}) = \lambda_{0Z_k}(t)\ e^{\beta_1 Z_1 + \ldots + \beta_{k-1} Z_{k-1} + \beta_{k+1} Z_{k+1} + \ldots + \beta_p Z_p}$$

Since we can estimate the survival function for any subgroup, we can use this to estimate the baseline survival function, $S_{0Z_k}(t)$, for each level of Z_k.

Then we compute $-\log S(t)$ for each level of Z_k, controlling for the other covariates in the model, and graphically check whether the log cumulative hazards are parallel across strata levels.

Example: PH assumption for `gender` (nursing home data):

- include `married` and `health` as covariates in a Cox PH model, but *stratify* by `gender`.

- calculate the baseline survival function for each level of the variable `gender` (i.e., males and females)

- plot the log-cumulative hazards for males and females and evaluate whether the lines (curves) are parallel

In the above example, we make the PH assumption for `married` and `health`, but not for `gender`.

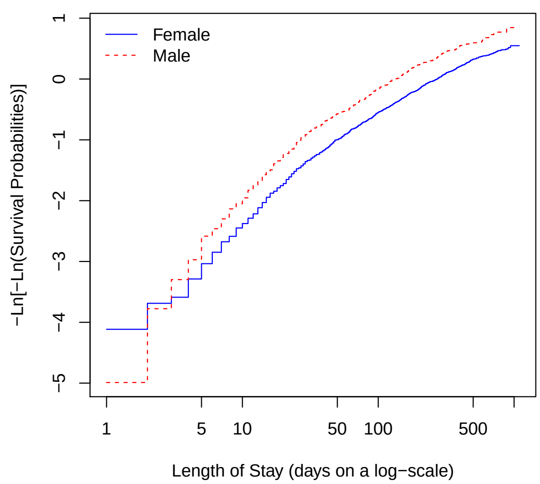
This is like getting a KM survival estimate for each gender without assuming PH, but is more flexible since we can control for other covariates.

We would repeat the stratification for each variable for which we wanted to check the PH assumption.

8.4.2 R Code for assesing PH within stratified models

```
fit.strat <- coxph( Surv(los,fail) ~ married + health
> + strata(gender),data = nurshome)
summary(fit.strat)

Call:
coxph(formula = Surv(los, fail) ~ married + health + strata(gender),
    data = nurshome)

  n= 1591, number of events= 1269

           coef exp(coef) se(coef)     z Pr(>|z|)
married 0.16101   1.17470  0.07720 2.086    0.037 *
health  0.16897   1.18408  0.03128 5.402  6.6e-08 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

        exp(coef) exp(-coef) lower .95 upper .95
married     1.175     0.8513     1.010     1.367
health      1.184     0.8445     1.114     1.259

Concordance= 0.55  (se = 0.011 )
Rsquare= 0.021   (max possible= 1 )
Likelihood ratio test= 33.97  on 2 df,   p=4.199e-08
Wald test            = 34.22  on 2 df,   p=3.702e-08
Score (logrank) test = 34.33  on 2 df,   p=3.518e-08

fitCoxgen2 = survfit(fit.strat)

plot(fitCoxgen2,mark.time = F,ylab = "-Ln[-Ln(Survival Probabilities)]",
     fun = "cloglog",
     xlab = "Length of Stay (days on a log-scale)",
     lty = 1:2,col = c("blue","red"))
legend("topleft",lty = 1:2,col = c("blue","red"),bty = "n",
       legend = c("Female","Male"))
```

Figure 8.4: Stratified model for gender controlling for marital status and health



8.5 Models with Time-dependent Interactions

Consider a PH model with two covariates Z_1 and Z_2. The standard PH model assumes

$$\lambda(t; Z) = \lambda_0(t)\ e^{\beta_1 Z_1 + \beta_2 Z_2}$$

However, if the log-hazards are not really parallel between the groups defined by Z_2, then you can add an interaction with time:

$$\lambda(t; Z) = \lambda_0(t)\ e^{\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_2 * t}$$

A test of the coefficient β_3 would be a test of the proportional hazards assumption for Z_2.

If β_3 is positive, then the hazard ratio would be increasing over time; if negative, then decreasing over time.

Changes in covariate status sometimes occur naturally during a study (ex. patient gets a kidney transplant), and are handled by introducing *time-dependent covariates*.

8.6 Comparing Cox PH survival to KM survival

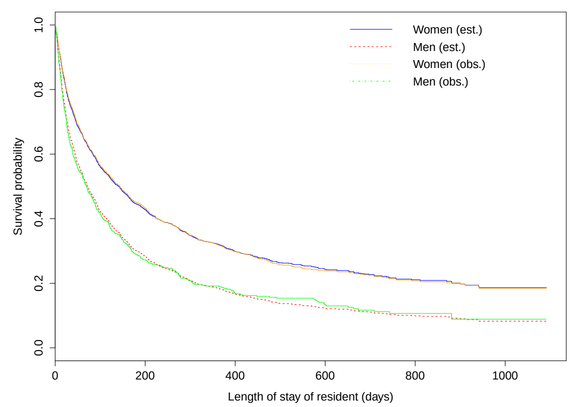
With R we simply overlay the plot of the Cox predicted survival on top of the KM survival.

The idea is that if the two survival curves coincide, then the Cox model reflects the data well.

The R code for the survival by gender is as follows:
Cox versus KM survival by gender:

The figure showing the predicted survival from the Cox model superimposed on the Kaplan Meier estimate of the survival is as follows:

Figure 8.5: Cox versus KM survival by gender



... or for a newly generated covariate (like `hlthsex`) which represents combined levels of more than one covariate.

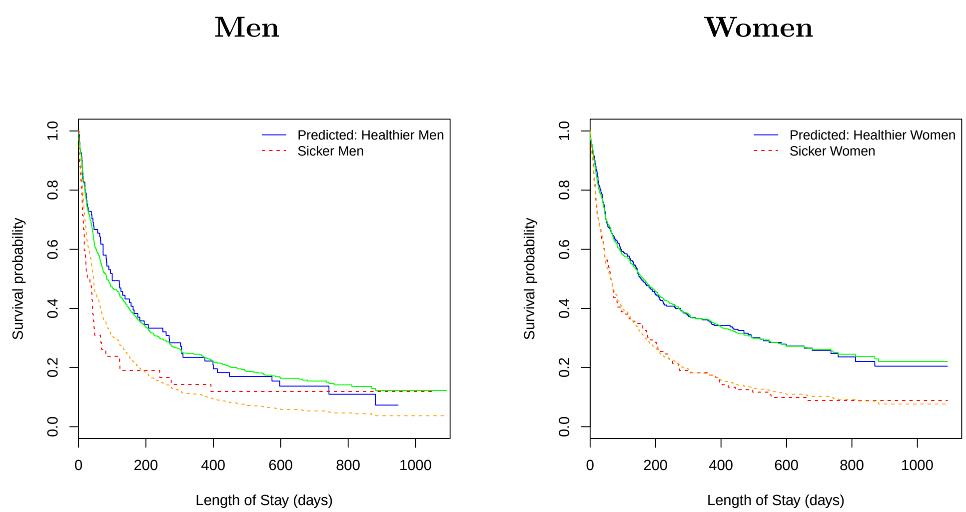
```
# Men
# Fit KM curves for males and females
fitKMgr.males = survfit( Surv(los,fail) ~ hlthsex,data = subset(nurshome2,
gender==1))

# Compare fitted with observed survival curves for gender
fitCoxgen.males = survfit( coxph( Surv(los,fail) ~ hlthsex,
                                   data = nurshome2),
                            newdata = data.frame(hlthsex =
                            as.factor(c("Healthier Men","Sicker Men"))))

plot(fitKMgr.males, mark.time = F,
```

```
     xlab = "Length of Stay (days)",ylab = "Survival probability",
     lty = 1:2,col = c("blue","red"))
# Add the raw curves
lines(fitCoxgen.males, lty = 1:2,col = c("green","orange"),mark.time = F)
legend("topright",bty = "n",lty = 1:4,col = c("blue","red","green","orange"),
       legend = c("Predicted: Healthier Men","Sicker Men"),ncol = 1,cex = 0.9)
```

Health and marital status:

The results are as follows: The comparison appears to be much better among

Figure 8.6: Comparison of Cox and KM survival curves for the combined health and marital status factors



women than among men.