computed using only the risk set at time y_j. A non-constant plot of observed S_j against y_j suggests this type of model failure.

These and other diagnostics for the proportional hazards model can be extended to time-dependent covariates.

Example 10.39 (PBC data) Primary biliary cirrhosis (PBC) is a chronic fatal disease of the liver, with an incidence of about 50 cases per million. Controlled clinic trials are hard to perform with very rare diseases, so the double-blinded randomized trial conducted at the Mayo Clinic from 1974–1984 is a valuable resource for liver specialists. A total of 424 patients were eligible for the trial, and the 312 who consented to take part were randomized to be treated either with the drug D-penicillamine or with a placebo. Although basic data are available on all 424 patients, we consider only these 312 individuals. Covariates available on each of them at recruitment include the demographic variables `sex` and `age`; clinical variables, namely presence or absence of `ascites`, `hepatomegaly`, `spiders`, and a ternary varable `edtrt` whose values $0, 1/2, 1$ indicate no, mild, and severe edema; and biochemical variables, namely levels of serum `bili`rubin (mg/dl), serum cholesterol (mg/dl), `al`bumin (gm/dl), urine copper (μg/day), `alk`aline `phos`phatase (U/ml), SGOT (U/ml), and triglycerides (mg/dl), `platelet` count (coded), `pro`thombin `time` (seconds), and the histologic stage of the disease (1–4). There are 28 missing values of serum cholesterol and 30 of triglycerides, and we ignore these covariates. Four missing values of platelets and two of urine copper were replaced by the medians of the remaining values; this should have little effect on the analysis. At the time at which the data considered here became available, 125 patients had died, with just 11 deaths not due to PBC, eight patients had been lost to follow-up, and 19 had undergone a liver transplant. As the response is time to death, these patients are regarded as censored.

Edema is the accumulation of fluids in body tissues.

The upper left panel of Figure 10.22 shows that estimated survivor functions for the patients with the drug and the placebo are very close, and it is no surprise that the log-rank statistic has value 0.1, insignificant when treated as χ^2_1. This is borne out by the estimated treatment effect of -0.057 (0.179) for a fit of the proportional hazards model with treatment effect only. Analysis stratified by sex gives an estimate of -0.045 (0.179). Neither differs significantly from zero. The corresponding baseline survival function estimates in the upper right panel of Figure 10.22 suggest no need to stratify.

Similar analyses for subgroups of the data and the corresponding log-rank statistics also show no significant treatment effects.

Having established that treatment has no effect on survival, we try constructing a model for prediction of survivor functions for new patients. This should be useful in assessing for whom liver transplant is a priority. The first step is to see which readily accessible covariates are highly predictive of survival. We exclude histologic stage, which requires a liver biopsy, and urine copper and SGOT, which are frequently unmeasured. The product-limit estimates and log rank statistics show strong dependence of failure on the other variables individually, so we fit a proportional hazards model

Variable	Estimate (SE)			
	Full	Reduced	Transformed	Final
age	0.028 (0.009)	0.030 (0.009)	0.033 (0.009)	0.041 (0.009)
alb	−0.97 (0.027)	−1.09 (0.24)	−3.06 (0.72)	−3.07 (0.72)
alkphos	0.015 (0.035)			
ascites	0.29 (0.31)			
bili	0.11 (0.02)	0.11 (0.02)	0.88 (0.10)	0.88 (0.10)
edtrt	0.69 (0.32)	0.77 (0.31)	0.79 (0.30)	0.69 (0.30)
hepmeg	0.49 (0.22)	0.50 (0.22)	0.25 (0.22)	
platelet	−0.61 (1.02)			
protime	0.24 (0.08)	0.25 (0.08)	3.01 (1.02)	3.57 (1.13)
sex	−0.48 (0.26)	−0.55 (0.25)		
spiders	0.29 (0.21)	0.30 (0.21)		

Table 10.23 Parameter estimates and standard errors for proportional hazards models fitted to the PBC data. The full fit is reduced by backwards elimination. In the last two columns log transformation is applied to `alb`, `bili`, and `protime`.

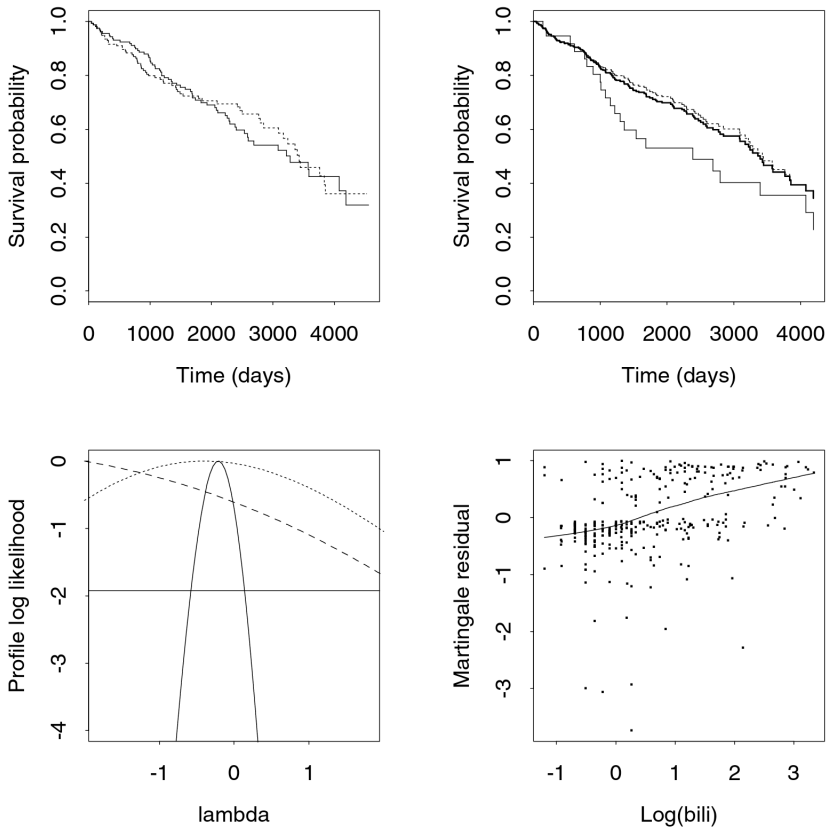

Figure 10.22 PBC data analysis (Fleming and Harrington, 1991). Top left: product-limit estimates for control (solid) and treatment (dots) groups. Top right: estimates of baseline survivor function for data stratified by sex, men (dots), women (solid). The heavy line shows the unstratified estimate. Lower left: profile likelihood for Box–Cox transformations of bilirubin (solid), albumin (dots), and prothrombin time (dashes); the horizontal line indicates 95% confidence limits for the transformation parameter. Lower right: martingale residuals from the model with terms `age`, `log(alb)`, `edtrt`, `log(protime)` against log bilirubin, and lowess smooth with $p = 2/3$.

with all but the excluded covariates. Table 10.23 suggests that serum bilirubin is most significant and that several other covariates can be dropped. Backward selection based on AIC leads to the reduced model in the table. The likelihood ratio statistic for comparison of the two models is 1.22, plainly insignificant. Dropping `sex` and

`spiders` also leads to a likelihood ratio statistic of 7.29, with significance level 0.20 when treated as χ^2_5. Bearing in mind the tendency of AIC to overfit, we now ignore these covariates.

To investigate whether transformation is worthwhile we apply the Box–Cox approach (Example 8.23) to `alb`, `bili`, and `protime`. The lower left panel of Figure 10.22 clearly indicates log transformation of `bili`, but not of the other variables. The need for transformation of `bili` can also be assessed through the plot of martingale residuals obtained when it is dropped, given in the lower right panel of the figure. Note the strong negative skewness of the residuals. The near-linearity of the lowess smooth shows the appropriateness of the transformation. The corresponding plot against `bili` itself is harder to read because the points are bunched towards zero. The plots for `alb` and `protime` are more ambiguous. If we take logs of all three variables, then the maximized log partial likelihood increases by 13.8 and `hepmeg` can be dropped; see Table 10.23.

A model with terms `age+log(alb)+log(bili)+edtrt+log(protime)` is medically plausible. As the disease progresses, the liver's ability to produce albumin decreases, leading to the negative coefficient for `alb`, while damage to the bile ducts reduces excretion of bilirubin and so increases its level in the body. Edema is often associated with the later stages of the disease, while prothrombin is decreased, leading to slower clotting of the blood. Finally and unsurprisingly, risk increases with age.

The upper panels of Figure 10.23 show deviance residuals plotted against age and prothrombin time. Inspection of those in the left panel lying outside the 0.01 and 0.99 normal quantiles reveals an error in the data coding; case 253 has residual -2.55 but his age should be 54.4 rather than 78.4. The right panel shows an unusually high prothrombin time of 17.1, which should have been 10.7. The estimates after these corrections are shown in the final column of Table 10.23.

The lower left panel of Figure 10.23 shows the scaled scores plotted against prothrombin time. There is some suggestion of non-proportionality, but it is too limited to suggest model failure. Such plots for the other variables cast no doubt on proportionality of hazards, and we accept the model.

To illustrate prediction, consider an individual with `age=60`, `alb=4`, `bili=1`, `edtrt=0`, and `protime=8`, for whom $x^{\mathrm{T}}\widehat{\beta} = -1.618$ and whose hazard is reduced by a factor $\exp(x^{\mathrm{T}}\widehat{\beta}) = 0.20$ compared to baseline. Setting `edtrt=1` and `bili=20` gives estimated risk scores of 0.4 and 2.8. The lower right panel of Figure 10.23 shows how the survivor functions then vary. The median estimated lifetime in each case can be found by solving for y the equation $\widehat{\mathcal{F}}_0(y)^{\exp(x^{\mathrm{T}}\widehat{\beta})} = 0.5$. ■

The proportional hazards model has been broadened in many directions. Suppose, for instance, that individuals move between states 1 and 2 and back again, baseline time-dependent transition rates $\gamma_{12}(y)$ and $\gamma_{21}(y)$ being modified to $\gamma_{12}(y)\xi_{12}(\beta; x)$ and $\gamma_{21}(y)\xi_{21}(\beta; x)$ for an individual with explanatory variables x. The partial likelihood for β is a product of terms corresponding to each of the observed transitions between states. For instance, the contribution from transition $1 \to 2$ at time y by an

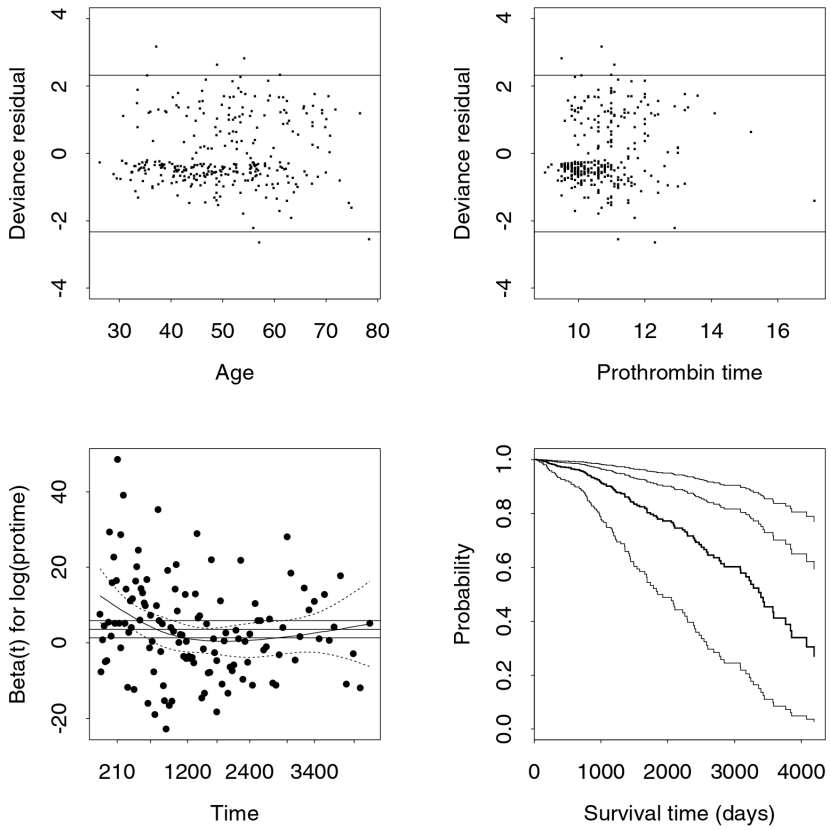

Figure 10.23 PBC data analysis. Upper panels: deviance residuals plotted against age and prothrombin time, with horizontal lines showing 0.01 and 0.99 standard normal quantiles. Lower left: scaled scores S_j^* plotted against prothrombin time, with lowess smooth and approximate 0.95 pointwise confidence bands (curved lines). Also shown are overall estimate and 0.95 confidence interval (horizontal lines). Lower right: baseline survivor function estimate (heavy), with predicted survivor functions for individuals with risk factors 0.2, 0.4, and 2.8 (top to bottom).

individual with covariates x_j is

$$\frac{\gamma_{12}(y)\xi_{12}(\beta;x_j)}{\sum \gamma_{12}(y)\xi_{12}(\beta;x_k)} = \frac{\xi_{12}(\beta;x_j)}{\sum \xi_{12}(\beta;x_k)},$$

the sum being over individuals in state 1 at time y. Individuals unobserved at y, or not in state 1, do not appear in the sum. Such extensions of partial likelihood enable inference for many types of partially observed and censored multi-state data, but details cannot be given here.

Counting processes and martingale residuals

This can be skipped at a first reading.

Consider a random variable Y with censoring indicator D and hazard function $h(y)$, and let

$$V(y) = I(Y \geq y), \quad N(y) = I(Y \leq y, D = 1),$$

be random variables that indicate whether Y is in view at time y, and whether failure has been observed by y. As $V(y)$ is left-continuous, its value at time y can be predicted the moment before, y^-, whereas the counting process $N(y)$ is right-continuous and so is not predictable. Let $\{\mathcal{H}_y : y \geq 0\}$ denote the history of the process up to time y. This is known as a filtration or increasing collection of sigma-algebras: $\mathcal{H}_x \subset \mathcal{H}_y$ for