Contents

9 Diagnostic Methods for Parametric Models (KM p. 409)

(Notes by Mark Eschmann, Updated by Sonish Lamsal)

- In the last sections, a variety of models for univariate survival data and several parametric models that can be used to study the effects of covariates on survival were presented.
- In this section, graphical checks of the appropriateness of these models is focused.
- Graphical checks of appropriateness is favored rather than formal statistical tests of lack of fit because these tests tend either to have low power for small-sample sizes or they always reject a given model for large samples.
- The graphical models discussed here serve as a model of rejecting clearly inappropriate models, not to "prove" that a particular parametric model is correct.

9.1 Checking adequacy of given model in a univariate setting

- The key idea is "to find a function of the cumulative hazard rate which is linear in some function of time."
- The basic plot is made by estimating the cumulative hazard eate by Nelson-aalen estimator.
- For example, consider the log logistic distribution.
  In this case $\widehat{H}(t) = \log(1 + \lambda t^{\alpha})$, thus we have that

$$\exp(\widehat{H}(t)) = 1 + \lambda t^{\alpha}$$
$$\Rightarrow \log\left[\exp(\widehat{H}(t)) - 1\right] = \log\lambda + \alpha\log t$$

  So, a plot of $ln\{exp[H(x)] - 1\}$ versus lnx should be approximately linear.

9.2 Other Models

Below is a table of other plots that can be made for some standard parametric models.

Model	Cumulative Hazard Rate	Plot
Log logistic	$\log(1 + \lambda t^{\alpha})$	$\log\left\{\exp\left(\widehat{H}(t)\right) - 1\right\}$ versus $\log(t)$
Exponential	λx	$\widehat{H}(t)$ versus t
Weibull	λx^{α}	$\log\left(\widehat{H}(t)\right)$ versus $\log(t)$
Log normal	$-\log\left\{1 - \Phi\left[\left(\log(t) - \mu\right)/\sigma\right]\right\}$	$\Phi^{-1}\left[1 - \exp\left(-\widehat{H}(t)\right)\right]$ versus $\log(t)$

Example 12.1 (continued)
Consider the Allo-Auto data set. Below are plots of the four diagnostic plots recommended above.

- In the first figure the Allo data appear to be nonlinear while the Auto data is roughly linear.
- For the other three figures, the book claims that they are roughly linear for both groups.

Note that the book generally chops off extreme x values. The probable justification is the small sample size in the tails although the book doesn't even mention that it is does this at all, and only a comparison of the figures that I generated versus figures 12.1-4 in the book shows that they did in fact do this.

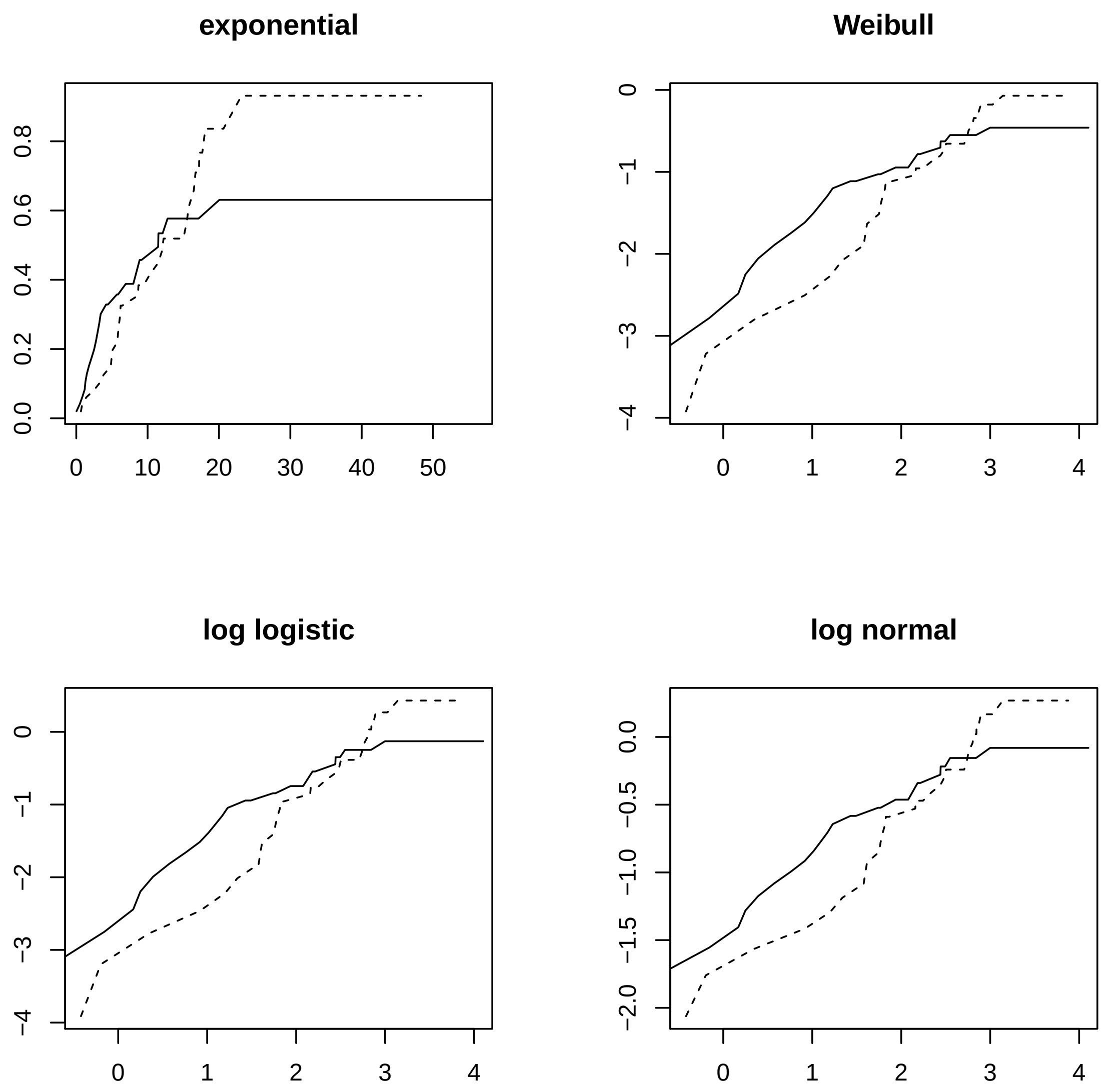


Figure 9.1: Hazard plots for all four parametric models in this section for allo (solid line) and auto (dashed line) transplant groups.

9.3 Checking appropriateness of accelerated failure-time model

- When comparing two groups, an alternative to the proportional hazards model is the accelerated failure-time model.
- A q-q plot can be used to determine the adequacy of this model
- The plot is based on the fact that, for the accelerated failure-time model,

$$S_1(t) = S_2(\theta t),$$

where S_0 and S_1 are the survival functions in the two groups and θ is the acceleration factor.

- Thus letting t_{0p} and t_{1p} be the p^{th} percentiles of the groups, we have the following relationship

$$t_{kp} = S_k^{-1}(1-p).$$

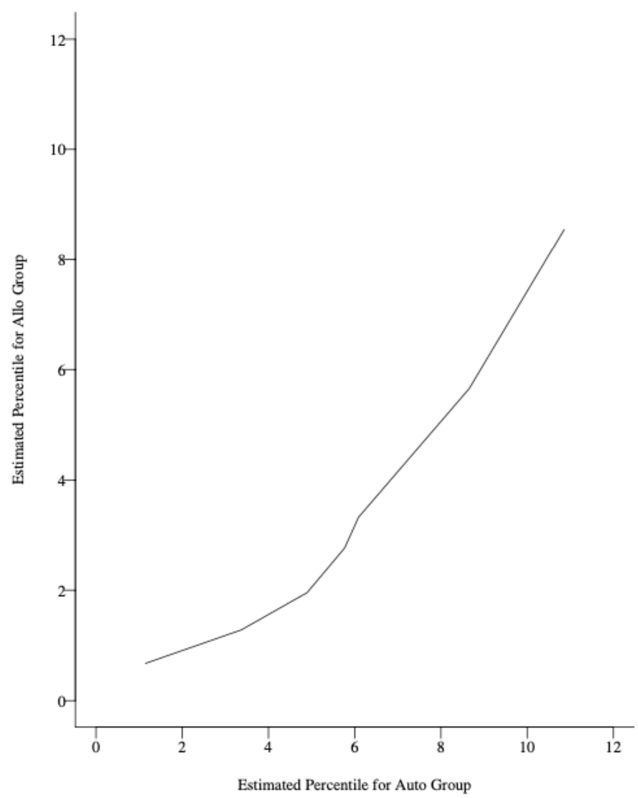
- Thus

$$S_0(t_{0p}) = 1 - p = S_1(t_{1p}) = S_0(\theta t_{1p}),$$

if the model holds.

- This can be checked by computing the Kaplan-Meier estimators of the two groups, estimating the percentiles t_{1p}, t_{2p} for various values of p, and then plotting the estimated percentiles in group 1 against the estimated percentiles in group 2. If the assumption holds, the graph should be a straight line with slope θ.

Example 12.1 (continued)
Again consider the Allo-Auto data set. A Kaplan-Meier estimate is calculated for each group and the percentiles $p = .05, .1, \cdots, .35$ are calculated for each group and plotted against each other below .



Note that only a portion of the range .05 to .35 is actually plotted above, but the portion that is plotted looks roughly linear. The figure appears to have a slope of about $\theta = 0.6$.

9.4 Parametric residuals

For the parametric regression problem, analogs of the residual plots described in Chapter 11 can be made with a redefinition of the various residuals to incorporate the parametric form of the baseline hazard rates.

9.4.1 Cox-Snell residuals

- The Cox–Snell residual is defined as $r_j = \widehat{H}(T_j|\mathbf{Z}_j)$, where $\widehat{H}$ is the fitted model. If the model fits the data, then the $r_j's$ should have a standard exponential distribution so that a hazard plot of r_j versus the Nelson-Aalen estimator of the cumulative hazard of the $r_j's$ should be a straight line with slope 1.

- The Cox Snell residuals of four parametric models are tabulated below.

Cox-Snell residuals	
Exponential	$\widehat{\lambda} t_i \exp\left\{\widehat{\beta}^T \mathbf{Z}_i\right\}$
Weibull	$\widehat{\lambda} \exp\left\{\widehat{\beta}^T \mathbf{Z}_i\right\} t_1^{\widehat{\alpha}}$
Log logistic	$\log \left[\frac{1}{1+\widehat{\lambda}\exp\left\{\widehat{\beta}^T \mathbf{Z}_i\right\} t_i^{\widehat{\alpha}}}\right]$
Log normal	$\log \left[1 - \Phi\left(\frac{\log T_i - \widehat{\mu} - \widehat{\gamma}^T \mathbf{Z}_i}{\widehat{\sigma}}\right)\right]$

- Equivalently we can analyze the data with the standardized residuals

$$s_j = \frac{\log T_j - \widehat{\mu} - \widehat{\gamma}^T \mathbf{Z}_j}{\widehat{\sigma}}$$

Example 12.2
In Figures 12.6–12.9, the cumulative hazard plots for the Cox–Snell residuals are shown for the exponential, Weibull, log logistic and log normal regression models for the laryngeal cancer data. We see from these plots that all four models give reasonable fits to the data, the best being the log normal and log logistic models.

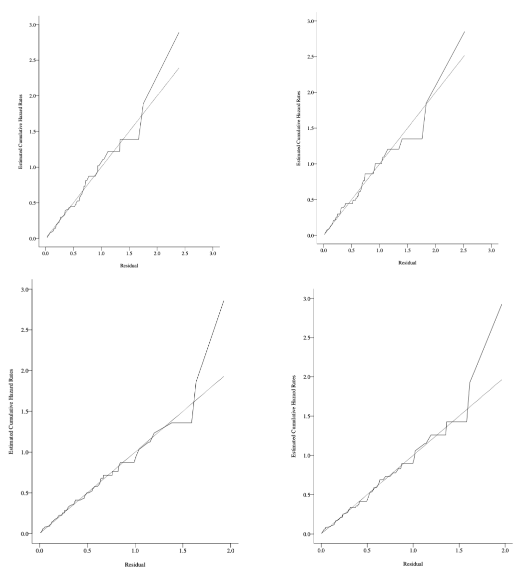


Figure 9.2: Cox–Snell residuals to assess the fit of (a) Exponential, (b) Weibull, (c) Log-logistic and (d) Log-normal regression model for the laryngeal cancer data set

9.4.2 Martingale and deviance residuals

- The Martingale residual is defined as

$$M_j = \delta_j - r_j.$$

- The deviance residual is defined as

$$D_j = \text{sign}[M_j]\left\{-2\left[M_j + \delta_j \log(\delta_j - M_j)\right]\right\}^{1/2}.$$

- As with the Cox model, the Martingale residual is an estimate of the excess number of deaths seen in the data, but not predicted by the model.
- The derivation of the Martingale in the Cox model does not hold here, but the name is the same because of the similar form
- The deviance residuals are an attempt to make the Martingale residuals symmetric about 0.
- Plots of either the Martingale or deviance residuals against time, observation number, or acceleration factor provides a check of model adequacy
- Basically, the use of the residuals is the same as in the previous chapter for the Cox model. Only their derivation has changed.

Example 12.2(Continued...)

- The fit of the log logistic regression model to the laryngeal cancer data using the deviance residuals is examined here.
- Below is a plot of the deviance residuals versus time on study.

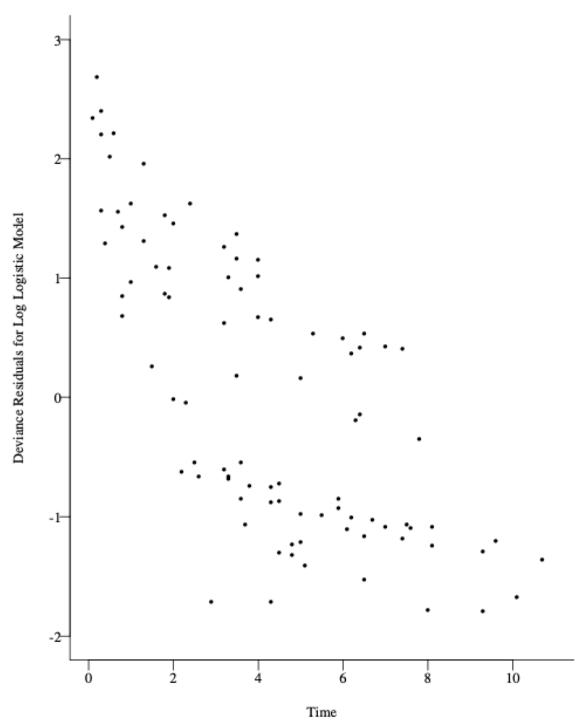


Figure 9.3: Deviance residuals from the log logistic regression model for laryngeal cancer patients

- Here, we see that the deviance residuals are quite large for small times and that they decrease with time.
- This suggests that the model underestimates the chance of dying for small t and overestimates this chance for large t.
- However, there are only a few outliers early, which may cause concern about the model.

9.5 Practical Note

Martingale and deviance residuals for these parametric models are available in S-Plus.