Contents

9 Diagnostic Methods for Parametric Models (KM p. 409)
(Notes by Mark Eschmann, Updated by Sonish Lamsal)
(Last Updated by Carson Slater (9.1- 9.2), Thomas Reinke (9.3) and Jacob Moore (9.4-9.5)) **2**
9.1 Checking Adequacy of Given Model in a Univariate Setting 2
9.2 Other Models 2
9.3 Checking appropriateness of accelerated failure-time model 3
9.4 Parametric residuals 5
9.4.1 Cox-Snell residuals 5
9.4.2 Martingale and deviance residuals 6
9.5 Practical Note 7

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

(Notes by Mark Eschmann, Updated by Sonish Lamsal)
(Last Updated by Carson Slater (9.1- 9.2), Thomas Reinke (9.3) and Jacob Moore (9.4-9.5))

- In the concluding sections, a range of models for univariate survival data and several parametric models were presented, which facilitate the investigation of the effects of covariates on survival outcomes.
- This subsection focuses on graphical checks of the appropriateness of these models, rather than relying on formal statistical tests of lack of fit.
- Graphical checks are preferred over formal statistical tests due to their superior performance in small-sample sizes, where they are more sensitive to deviations from model assumptions. Formal statistical tests in this context are criticized due to low power.
- The graphical models discussed in this context serve as a paradigm for identifying and rejecting clearly inappropriate models, rather than providing conclusive evidence for the correctness of a particular parametric model.

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 rate 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

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

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

9.2 Other Models

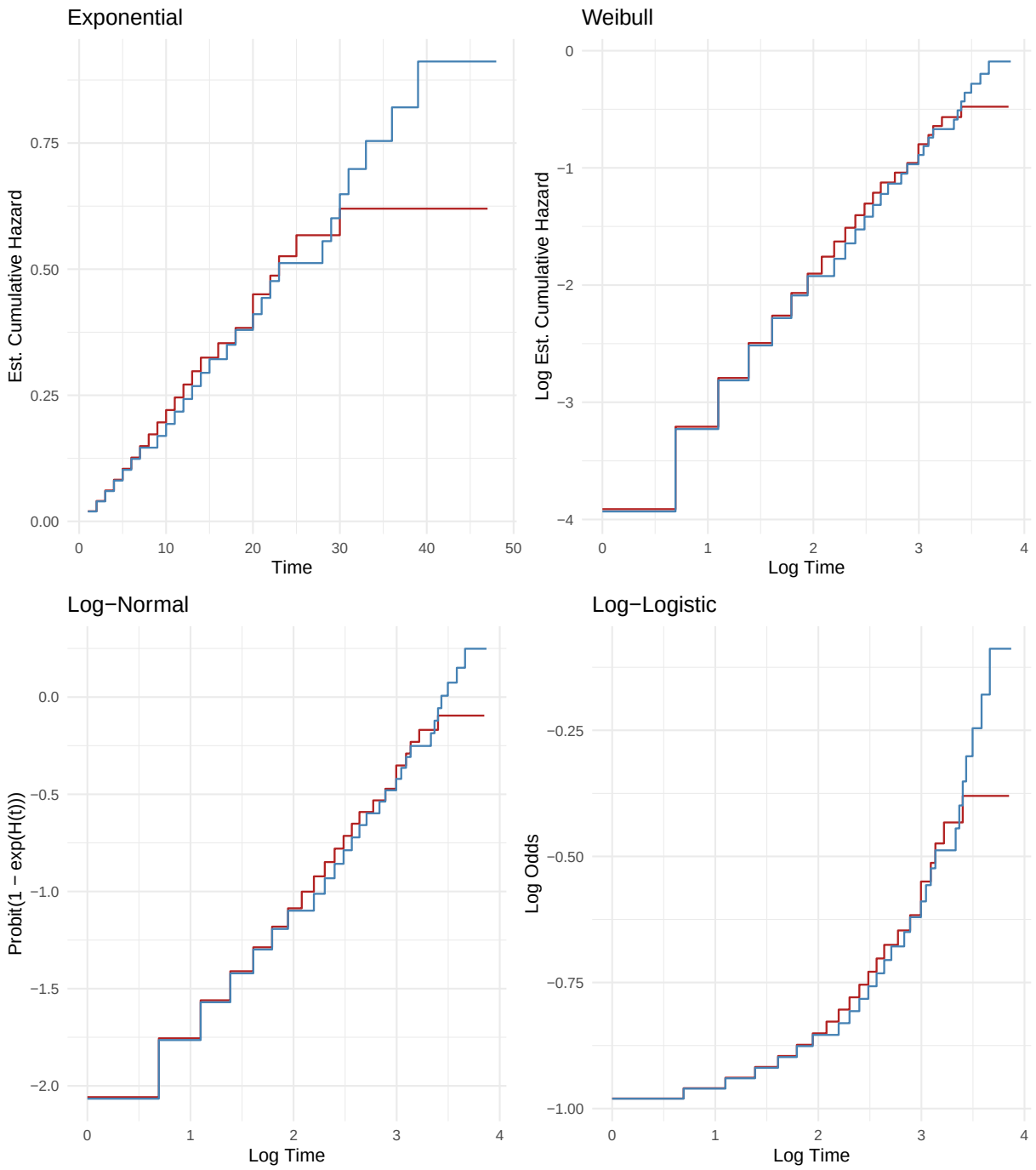
Below is a table of other plots that can be made for some standard parametric models.

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

Example: Waiting Time to Transplant (KM pp. 411-413):
Consider the data presented in Section 1.10 of KM comparing allogeneic (Allo) bone marrow transplants from an HLA-matched sibling or an autogeneic (Auto) bone marrow transplant for patients with either Hodgkin's disease (HL) or non-Hodgkin's lymphoma (NHL). 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.[1]

[1] Note that the book generally chops off extreme t values. The probable justification is the small sample size in the tails, although the book doesn't even mention that it 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.



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 QQ plot is 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.

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

$$t_{kp} = S_k^{-1}(1-p),\ k = 0, 1$$

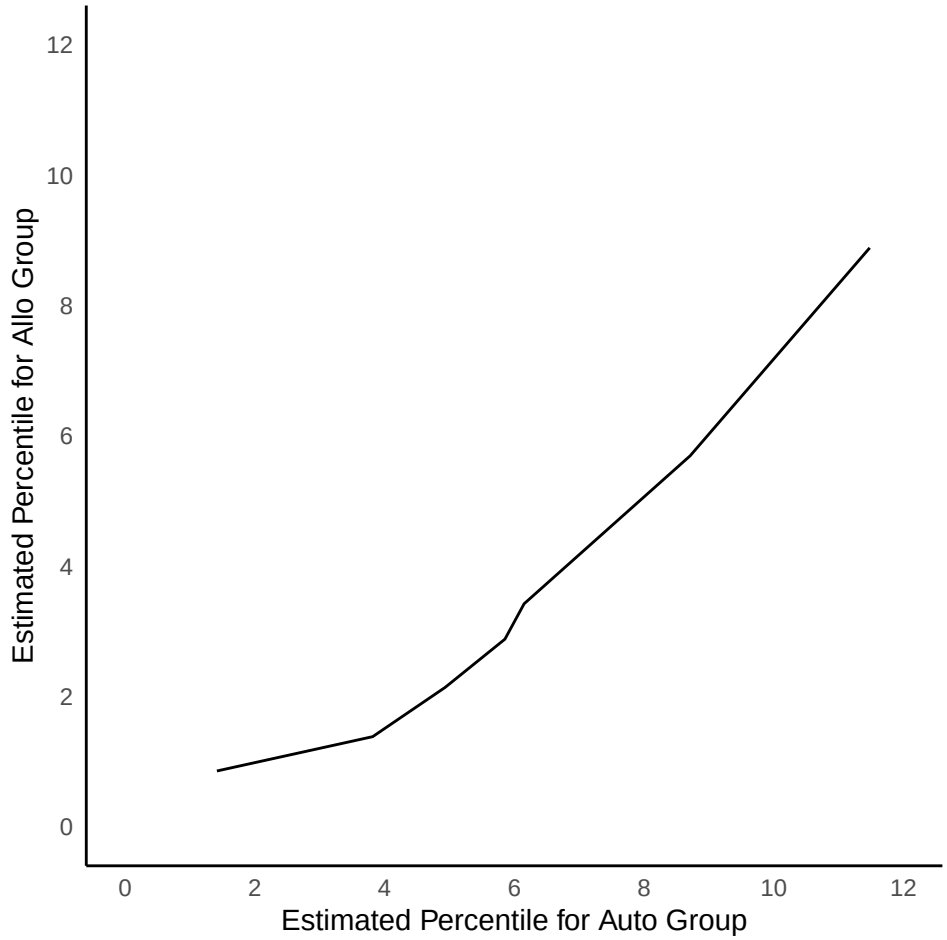
- Thus

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

  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. If the curve is linear, then a crude estimate of the acceleration factor q is given by the slop of the line.

Example 12.1 (continued)
Again consider the Allo-Auto data set with allogenic transplants (group 0) and autologous translants (group 1). 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 the range of percentiles .05 to .35 is actually plotted above, but the portion that is plotted looks roughly linear. The curve appears to have a slope q between 0.6 and 0.8, which is our crude estimate for the acceleration factor θ. [2]

[2]Klein & Moeschberger do not mention why only percentiles from .05–.35 are chosen. A likely reason is that for the Allogenic group, the estimated percentile values are NA after $p > .45$

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.

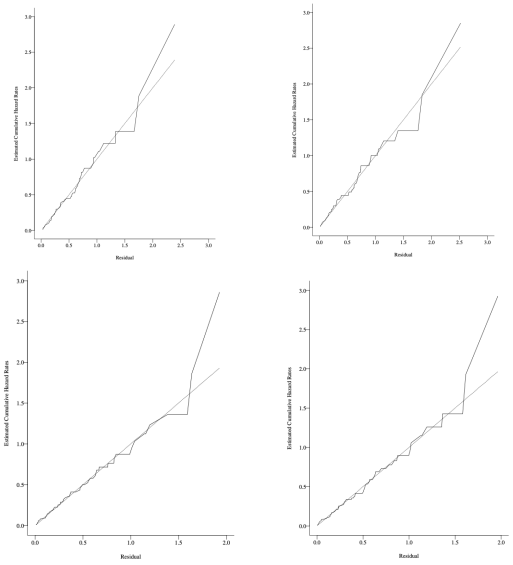


Figure 9.1: Cox–Snell residuals to assess the fit of (a) Exponential, (b) Weibull, (c) Log-logistic and (d) Log-normal regression model for the laryngeal cancer data set

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.

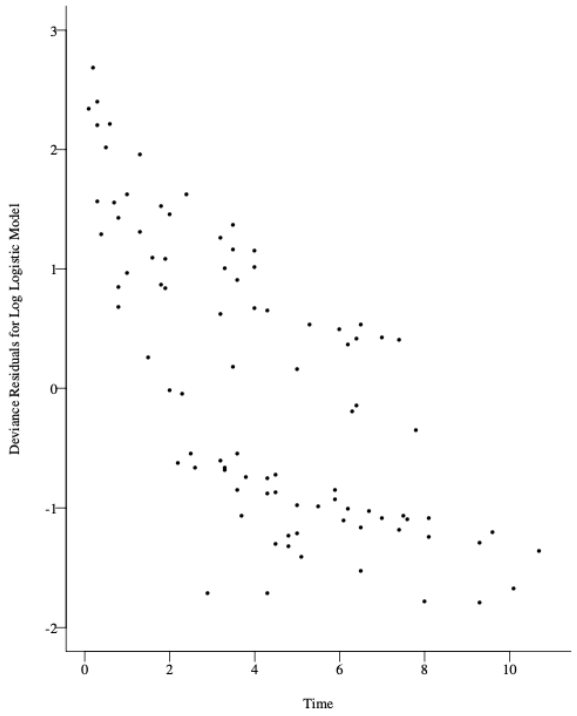


Figure 9.2: Deviance residuals from the log logistic regression model for laryngeal cancer patients

- 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

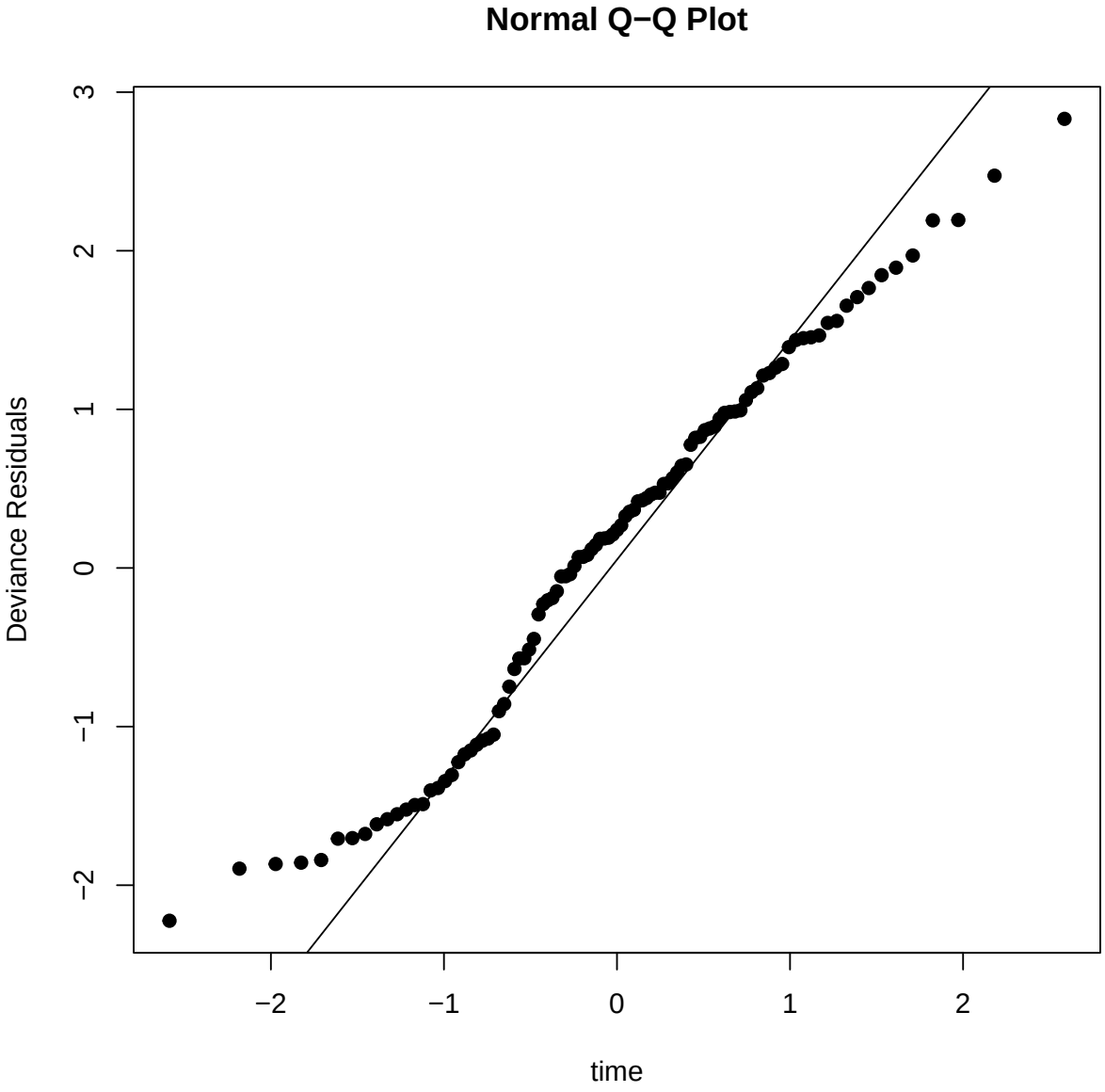
Martingale and deviance residuals for these parametric models are available in R with the 'survival' package.
Example Heart data

Survival of patients on the waiting list for the Stanford heart transplant program. (Found as the example data in R). Covariates include age, an indicator for if the patient has had a prior bypass surgery (surgery), an indicator for if they've had a heart transplant (transplant), and interaction between age and transplant as well as surgery and transplant. ($n = 172$)

```
library(survival)
data(heart)

fit <- coxph(Surv(start, stop, event) ~ (age + surgery) * transplant,
             data=heart)
mresid <- resid(fit, collapse=heart$id) #Martingale Residuals
dresid <- resid(fit, collapse=heart$id, type = "deviance") #Deviance Residuals
```

```
qqnorm(dresid, xlab = "time", ylab = "Deviance Residuals", pch = 19)
qqline(dresid)
```


Normal Q−Q Plot
Deviance Residuals
time
3
2
1
0
−1
−2
−2
−1
0
1
2