Actuarial Science 4823 Final Project: Recurrence Times of Colon Cancer under Levamisole and 5-FU Chemotherapy Agents

Daniel Grant
Ruobing Wei
Sana Mungroo
December 18th, 2019

Contents

Objectives

In 1989, Laurie et al. conducted a study to investigate the effectiveness of 2 drugs (Levamisole) and (Fluorouracil) as a new adjuvant therapy for large-bowel carcinoma. As such, patients in the study were those who had, prior to the study, stage B or C cancer removed surgically. One of three interventions was used on the set of patients: observation only, levamisole treatment alone, or a combination of levamisole plus fluorouracil treatments[3].

In 1995 Moertel et al. wished to further investigate the efficacy of these drugs on patients with stage III (stage C) disease specifically. Ultimately, they wished to have these drugs be considered as valid treatments to reduce recurrences and deaths caused by this cancer[6].

The objective of this analysis is more limited than what Laurie and Moertel had investigated. Specifically, we attempt to ascertain whether there does exist a significant difference between the adjuvant therapies in regards to recurrence times, as well as determining which covariates seem to be most associated with lower (or higher) recurrence times. Gaining insight into the latter has the potential to allow researchers to better understand how the cancer spreads and what aspects of the disease their treatments should target.

Data

This data consists of observations corresponding to 929 patients, with each patient being represented with 2 rows in the data, the first being the row corresponding to the cancer's recurrence time, and the second being the row corresponding to the death time. If recurrence (and thus death) was not observed by the end of the study, the observation was considered (type I right) censored. This was likely due to the fact that it would become infeasible to continue to follow patients much beyond the 5-6 years that patients were followed in reality.

Attributes of this data set included patient ID, treatment given, sex, age, as well as various attributes of the cancer that were determined on or prior to the patient's entrance into the study. This data also contained 41 observations for which there was at least 1 missing element, and these observations were thus removed, leaving a total of 888 rows for future analysis. A description of the variables of this data set, along with the first 9 rows is provided below.

Variable	Description
id	Patient ID
study	1 for all patients
rx	Treatment:
	Obs(ervation)
	Lev(amisole)
	Lev(amisole)+5-FU (Fluorouracil)
sex	Patient sex (1=male)
age	Age (in years)
obstruct	Obstruction of colon by tumor (1=yes)
perfor	Perforation of colon (1=yes)
adhere	Adherence to nearby organs (1=yes)
nodes	Number of lymph nodes with detectable cancer
status	Censoring status (0=censored)
differ	Differentiation of tumor (1=well, 2=moderate, 3=poor)
extent	Extent of local spread (1=submucosa, 2=muscle, 3=serosa, 4=contiguous structures)
surg	Time from surgery to registration (0=short, 1=long)
node4	More than 4 positive lymph nodes (1=yes)
time	Days until event or censoring
etype	Event type (1=recurrence, 2=death)

Table 1: Descriptions of variables contained in colon

Table 2: colon

| id	study	rx	sex	age	obstruct	perfor	adhere	nodes	status	differ	extent	surg	node4	time	etype
1	1	Lev+5FU	1	43	0	0	0	5	1	2	3	0	1	1521	2
1	1	Lev+5FU	1	43	0	0	0	5	1	2	3	0	1	968	1
2	1	Lev+5FU	1	63	0	0	0	1	0	2	3	0	0	3087	2
2	1	Lev+5FU	1	63	0	0	0	1	0	2	3	0	0	3087	1
3	1	Obs	0	71	0	0	1	7	1	2	2	0	1	963	2
3	1	Obs	0	71	0	0	1	7	1	2	2	0	1	542	1
4	1	Lev+5FU	0	66	1	0	0	6	1	2	3	1	1	293	2
4	1	Lev+5FU	0	66	1	0	0	6	1	2	3	1	1	245	1
5	1	Obs	1	69	0	0	0	22	1	2	3	1	1	659	2

Because the focus of this analysis is solely on recurrence times, only rows with etype=2 were used in the analysis. Note also that some censoring occurred 3087 days (over 8 years) after entering the study. While it would be interesting to observe patients until their natural deaths, monitoring events much beyond 8 years is not likely to be feasible in most situations.

Preliminary Analysis

Prior to model selection, it may be of interest to run a preliminary analysis using nonparametric methods to gain insight as to whether certain covariates may have different effects on recurrence time.

Of primary interest is whether different treatments in `rx` will have these differing effects. By constructing Kaplan Meier estimates of the survival probability over time for each therapy, it appears as if the administering levamisole exclusively was no more effective than merely observing patients. However, patients who were administered levamisole plus fluorouracil appear to have much higher probabilities of their cancer not having recurred by a given time. Note that black "+" symbols indicate that censoring occured in that observation, which in this dataset almost exclusively means that the patient's cancer did not recurr within the length of the study. In this way, it is of note is that these survival curves appear to level-off when time becomes large, this suggests that these patients are "cured" of the cancer after time increases.

Kaplan Meier Survival Probability Estimates
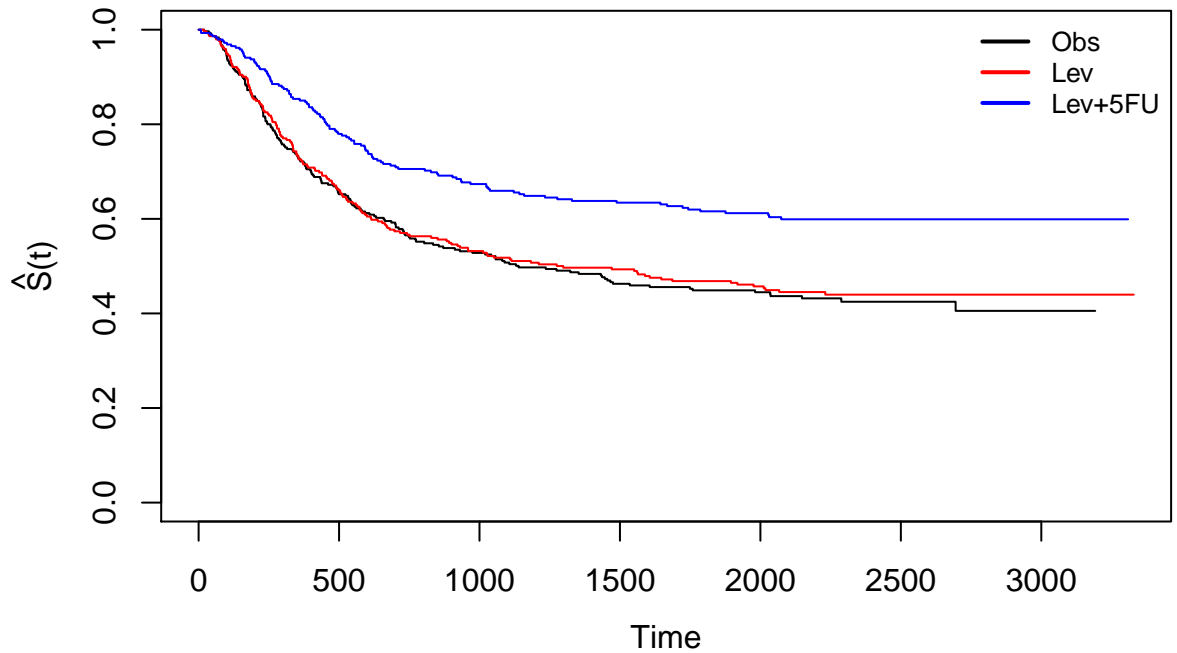
As such, `survdiff()` was used to perform the Log Rank test on these different treatment groups. This test works by utilizing the fact that under the null hypothesis $H_0 : S_{Obs}(t) = S_{Lev}(t) = S_{Lev+5FU}(t)$, the sum of squared differences between observed and expected events for each treatment, when divided by their respective excpected number of events will approximately follow a χ^2_2.

Table 3: Log-Rank test of differences between `rx`

|	N	Observed	Expected	$(O-E)^2/E$	$(O-E)^2/V$
rx=Obs	305	172	145	5.13	7.60
rx=Lev	294	161	141	2.84	4.15
rx=Lev+5FU	289	113	160	13.93	21.80
$\chi^2 = 22$ on 2 df	$p = 2 \times 10^{-5}$				

This test suggests that at least one of the adjuvant therapies differs significantly from the others, which

implies that at least one of the treatments is expected to differ significantly from mere observation, and this therapy may be the levamisole and fluorouracil therapy.

AFT Model

A natural question at this stage in the analysis was whether an accelerated failure time (AFT) or proportional hazard (PH) model was most appropriate modeling recurrence times. The difference between these models is that, loosely, an AFT model may be interpreted as modeling changes in the time (or log time) to an event (in this case recurrence) under different covariates, while a PH model seeks to describe the effect these covariates have on the instantaneous risk (the hazard) of recurrence.

In order to gain some intuition as to which model might be more appropriate, Kaplan-Meier estimates of the survival function (denoted $\hat{S}_k(t)$) were constructed under models using only the kth covariate. Since there are 11 covariates in this data set, 11 different single-covariate models were constructed. If $\log(-\log \hat{S}_k(t))$ is plotted against $\log t$, ideally all 11 plots would show the same general trend: either all vertical trending parallel lines or all horizontally trending parallel lines. If these plots have a horizontal trend, this indicates that an AFT model may be most appropriate, while if the plots have a vertical trend, a PH model may be most appropriate.
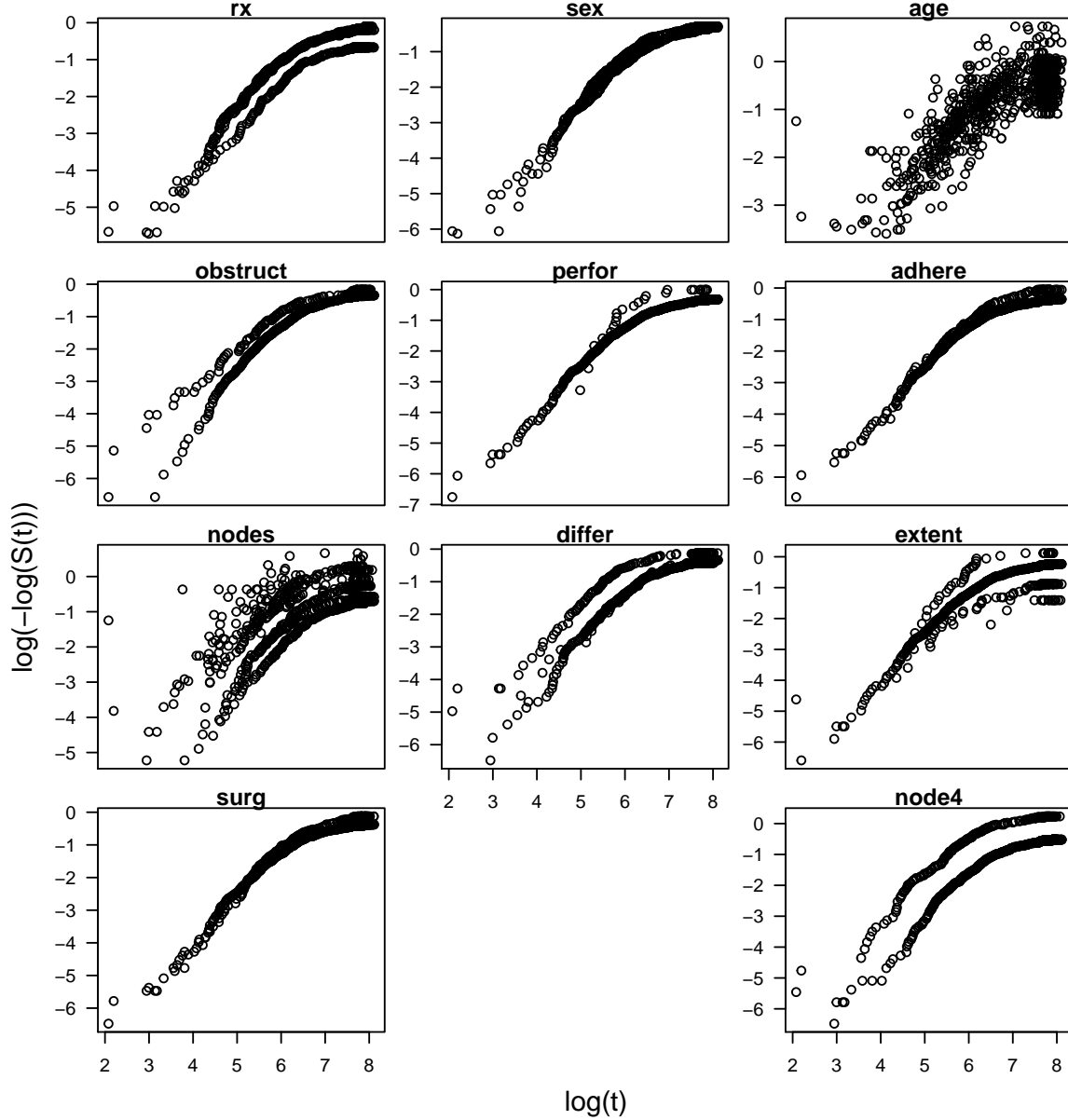
In the plots above, neither a vertical nor horizontal trend is apparent. Despite not following a strictly linear pattern, it may be wise to choose a Weibull lifetime model. This is convenient in part because Weibull models are both AFT and PH models, so making the incorrect choice with regard to assuming accelerated failure times or proportional hazards will not cause problems here. As an alternative to the aforementioned parametric AFT (and PH) model, a semi-parametric Cox PH model will be constructed, since this model leaves the baseline hazard $h_0(t)$ unspecified, and we may later estimate it solely using the known data if we wished. In essence, this method allows for the analysis of differences between levels of covariates, without needing to examine the total risk being incurred by a patient. This semi-parametric model would not be optimal for making future predictions on new data, but since this study is concerned with data already known and obtained, this is not an issue.

Using `survreg()`, an AFT model was constructed using with the Weibull distribution specified, and all 11 covariates contained in the model. As suggested by Lawless(p.273), this Weibull assumption may also lead to an alternative PH interpretation with the baseline hazard being specified parametrically as that of a

Weibull[5]. The estimated coefficients associated with the covariates, $\hat{\beta}_{PH}$, would then be given as

$$\hat{\beta}_{PH} = -\frac{1}{\hat{b}}\hat{\beta}_{AFT}$$

where $\hat{\beta}_{AFT}$ are the AFT model estimates, and $\hat{b}$ is the estimated scale parameter associated with the AFT model. Note that this PH interpretation is not the same as that given via `coxph()`. This is because the Weibull hazard assumption means that our PH model is fully parametric, while `coxph()` leaves the hazard function unspecified, meaning that it will be used to construct the semi-parametric model second model that is being proposed. A summary of the AFT model is provided below,

Table 4: Summary of full AFT model

	Est. Value	Std. Error	z	p	
(Intercept)	9.303	0.779	11.939	2.0×10^{-16}	***
rxLev	0.080	0.152	0.524	0.298	
rxLev+5FU	0.75437	0.16771	4.498	6.9×10^{-6}	***
sex1	0.158	0.131	1.200	0.230	
age	0.003	0.006	0.532	0.595	
obstruct1	-0.229	0.164	-1.400	0.162	
perfor1	-0.314	0.351	-0.896	0.370	
adhere1	-0.251	0.180	-1.395	0.163	
nodes	-0.061	0.021	-2.958	0.003	***
differ2	0.058	0.226	0.257	0.797	
differ3	-0.261	0.265	-0.984	0.325	
extent2	-0.414	0.727	-0.569	0.569	
extent3	-1.155	0.694	-1.665	0.096	*
extent4	-1.686	0.750	-2.249	0.025	**
surg1	-0.346	0.143	-2.418	0.016	**
node4	-0.812	0.194	-4.178	0.00003	***
Log(scale)	0.312	0.0412	7.584	3.4×10^{-14}	***
Scale = 1.366					
Observations	888				
Log Lik(model)	-3862.43	Log Lik(intercept)	-3938.1		
χ^2 (df = 15)	151.42***				

Note: *p<0.1; **p<0.05; ***p<0.01

    We may briefly consider the aforementioned PH interpretation of this data under the Weibull assumption by utilizing `flexsurvreg()` from the package `flexsurv` which, unlike the traditional function `survreg()`, allows for specification of the baseline hazard using `dist="weibullPH"`.

Table 5: Estimated β_j for alternative PH interpretation

	$Est.Value$
rxLev	-0.058
rxLev+5FU	-0.552
sex1	-0.115
age	-0.002
obstruct1	0.168
perfor1	0.230
adhere1	0.184
nodes	0.045
differ2	-0.042
differ3	0.191
extent2	0.303
extent3	0.845
extent4	1.234
surg1	0.253
node4	0.594
shape	-0.312
scale	-6.809

It is possible to check using Table 4 and Table 5 that the estimated coefficients in Table 11 do indeed follow $\hat{\beta}_{PH} = -\frac{1}{b}\hat{\beta}_{AFT}$.

From Table 4, it can be seen that `rx=Lev+5FU`,`nodes`,`extent=4`,`surg=1`, and `node4=1` are significant at the $\alpha = 0.05$ level. Despite a lack of evidence suggesting the estimated coefficient associated with `extent=2` and `extent=3` differs from 0, this variable must remain in future models, since other levels of this categorical variable are significant and must remain.

`surg=1` being significant suggests that longer times from surgery to registration were significantly associated with faster recurrence of the cancer, and having more than 4 cancer-positive lymph nodes were also assocaited with a faster recurrence than having 4 or fewer.

Most importantly, there is not enough evidence to suggest that `rxLev` is associated with a difference in recurrence times, since the p-value associated with its effect estimate is 0.6. Notably, however, the effect associated with the treatment of a combination of Levamisole and Fluorouracil was 6.9×10^{-6}, meaning that there is evidence to support an increase in time until recurrence under treatment with this drug. Furthermore, this effect was estimated as increasing log time until recurrence time by 0.754 compared to the patients who were merely observed. This effect translates to a predicted time until recurrence of 2.126 times that of an observed patient, all other factors remaining the same.

In order to remove hypothesized non-significant predictors, the likelihood ratio test was performed with the null hypothesis that the β coefficients associated with `sex`,`age`,`obstruct`,`perfor`, and `differ` are in fact 0. A reduced model was fitted assuming these coefficients were in fact 0, and the log likelihood of each model is shown below

$$2(\ell_{Full} - \ell_{Reduced}) \sim \chi_5^2$$
$$2(-3862.4 - (-3865.3)) = 5.8 < 11.07 = \chi_5^{2\ 0.95}$$

This suggests that there is not significant evidence to support the fact that these coefficients are all nonzero, as was anticipated.

Table 6: Reduced AFT model

	Value	Std. Error	z	p	
(Intercept)	9.555	0.695	13.751	2.0×10^{-16}	***
rxLev	0.080	0.152	0.525	0.5995	
rxLev+5FU	0.736	0.168	4.387	1.1×10^{-5}	***
obstruct1	-0.259	0.162	-1.603	0.1089	
adhere1	-0.305	0.175	-1.742	0.0815	*
nodes	-0.064	0.021	-3.054	0.0023	***
extent2	-0.349	0.728	-0.479	0.6317	
extent3	-1.127	0.695	-1.622	0.1048	
extent4	-1.695	0.750	-2.260	0.0238	**
surg1	-0.346	0.143	-2.420	0.0155	**
node4	-0.824	0.195	-4.214	2.5×10^{-5}	***
Log(scale)	0.3151	0.041	7.639	2.2×10^{-14}	***
Scale $=$	1.37				
Observations	888				
Log Lik(model)	-3,865.326	Log Lik(intercept)	-3,938.140		
χ^2 (df $=$ 10)	145.629***				

Note: *p<0.1; **p<0.05; ***p<0.01

Semi-parametric Cox PH Model

With regard to the second proposed model, this model was fit through `coxPH`, and coefficient estimates are displayed below.

Table 7: Estimated β_j for semi$-$parametric Cox PH

	coef	exp(coef)	Std. Error	z	p	
rxLev	-0.038	0.963	0.112	-0.338	0.735	
rxLev+5FU	-0.508	0.602	0.122	-4.164	0.00003	***
sex1	-0.129	0.879	0.096	-1.343	0.179	
age	-0.003	0.997	0.004	-0.632	0.527	
obstruct1	0.192	1.211	0.120	1.602	0.109	
perfor1	0.202	1.224	0.257	0.787	0.431	
adhere1	0.163	1.177	0.131	1.239	0.215	
nodes	0.038	1.038	0.015	2.514	0.012	**
differ2	-0.038	0.962	0.165	-0.233	0.816	
differ3	0.229	1.258	0.194	1.184	0.237	
extent2	0.284	1.329	0.532	0.535	0.593	
extent3	0.793	2.209	0.507	1.564	0.118	
extent4	1.206	3.340	0.547	2.203	0.028	**
surg1	0.236	1.266	0.104	2.265	0.024	**
node4	0.595	1.814	0.141	4.223	0.00002	***
Observations	888					
Log Lik(model)	-2,811.299		Log Lik(intercept)	-2,877.133		
Wald Test (df = 15)	143.700***					
LR Test (df = 15)	138.722***					
Score Test (df = 15)	153.998***					

Note: *p<0.1; **p<0.05; ***p<0.01

Again using the likelihood ratio test, it was necessary to test if `sex`, `age`, `obstruct`, `perfor`, and `differ` could be removed from the model. The log likelihood of the full model and the reduced model were used to perform the test below,

$$2(\ell_{Full} - \ell_{Reduced}) \sim \chi_5^2$$
$$2(-2807.772 - (-2811.299)) = 7.04 < 11.07 = \chi_5^{2\ 0.95}$$

Again there is not enough evidence to suggest these coefficients are nonzero, and thus the reduced model is acceptable in that regard.

This PH model suggests the same variables are significant as were suggested in the reduced AFT model.

Table 8: Estimated β_j for reduced semi$-$parametric Cox PH

	coef	exp(coef)	Std. Error	z	p	
rxLev	-0.040	0.961	0.111	-0.358	0.721	
rxLev+5FU	-0.492	0.612	0.122	-4.039	0.0001	***
obstruct1	0.211	1.235	0.118	1.795	0.073	*
adhere1	0.200	1.222	0.128	1.571	0.116	
nodes	0.040	1.040	0.015	2.628	0.009	***
extent2	0.228	1.256	0.531	0.430	0.667	
extent3	0.763	2.145	0.506	1.509	0.131	
extent4	1.206	3.339	0.546	2.209	0.027	**
surg1	0.232	1.261	0.104	2.234	0.026	**
node4	0.604	1.829	0.141	4.271	0.00002	***
Observations	888					
Log Lik(model)	-2,811.299		Log Lik(intercept)	-2,877.133		
Wald Test (df = 10)	137.060***					
LR Test (df = 10)	131.668***					
Score Test (df = 10)	146.803***					

Note: *p<0.1; **p<0.05; ***p<0.01

Analysis

Examining the reduced AFT model described in Table 5, the estimated β_j should be interpreted as the effect that increasing the covariate by 1 unit (or changing the level relative to the base case for categorical covariates) has on log time to recurrence (which we have assumed is log Weibull i.e. extreme value). This can be translated into an effect on recurrence time directly by $\exp(\hat{\beta}_j)$. These effects work multiplicatively with time, to decelerate (>1) or accelerate (<1) the time to recurrence by a factor of $\exp(\hat{\beta}_j)$.

Table 9: Deceleration factor for all covariates

	Time to Recurrence Factor
rxLev	1.083
rxLev+5FU	2.087
obstruct1	0.772
perfor1	0.730
adhere1	0.737
nodes	0.938
extent2	0.706
extent3	0.324
extent4	0.184
surg1	0.708
node4	0.439

For example, this model suggests that relative to being receiving no adjuvant therapy and merely be observed, being given levamisole and fluorouracil is associated with a 2.087 times increase in time to recurrence (i.e. that being given Lev+5FU is associated with recurrence at $2.087^{-1} = 0.479 = 47.9\%$ the speed of `rx=Obs` recurrence). Conversely, in the case of severe spread of the cancer (`extent=4`), recurrence is expected to occur at $0.184^{-1} = 5.43 = 543.0\%$ the speed of a cancer with `extent=1`, all else remaining the same.

The interpretation of coefficients in the case of the Cox PH model differs from that of the AFT model. This model is of the form

$$h(t|x) = h_0(t)e^{\beta_1 x_1 + \cdots + \beta_{10} x_{10}}$$

Where $h_0(t)$ remains unspecified. This means that when increasing a covariate x_i by 1 unit and leaving all other covariates the same, the hazard ratio becomes

$$\frac{h(t|x_1, \ldots, x_i + 1, \ldots x_{10})}{h(t|x_1, \ldots, x_i, \ldots x_{10})} = \frac{h_0(t)e^{\beta_1 x_1 + \cdots + \beta_i(x_i+1) + \ldots \beta_{10} x_{10}}}{h_0(t)e^{\beta_1 x_1 + \cdots + \beta_i x_i + \ldots \beta_{10} x_{10}}} = e^{\beta_i}$$

We may either take the log of this to write in terms of β_i, or directly interpret $e^{\beta_i} - 1$ as the percent change in hazard given a 1 unit increase (or increase relative to the baseline) of covariate i. This information is derived using the second column of Table 7

Table 10: Percentage change in hazard for 1 unit change (or change relative to baseline) in covariate

	Change in Hazard
rxLev	-3.89%
rxLev+5FU	-38.85%
obstruct1	23.54%
adhere1	22.19%
nodes	4.03%
extent2	25.61%
extent3	114.54%
extent4	233.92%
surg1	26.14%
node4	82.92%

Applications

In the case of either model, it may be of interest what a certain patient's recurrence time being larger than a certain time is, given their diagnosis. For example, the median recurrence time under the AFT model would be given, using the fact that the median of a standard extreme value distribution (log Weibull), denoted $w_{0.5}$ is

$$e^{-e^{w_{0.5}}} = 0.5 \implies w_{0.5} = \log(-\log(0.5)) = -0.367$$

Then under the AFT model, Y has an extreme value distribution with scale and location $(\hat{\beta}_j)$ parameters given directly through the tables shown above. Thus the median log-recurrence time is

$$y_{0.5} = \beta_0 + \cdots + \beta_{10} x_{10} + w_{0.5} b$$

replacing these parameters with their estimates, and assuming the patient of interest was: treated with levamisole and fluorouracil, had obstruction, adherance, 5 nodes, and was in extent 4 with a short time from surgery to regististration,

$$y_{0.5} = 9.555 + 0.736 - 0.259 - 0.3052 - 0.064(5) - 1.695 - 0.823 - 0.367(1.370) = 6.386$$
$$\implies t_{0.5} = \exp(6.386) = 981.223$$

Or just over 1.5 years until median recurrence time.

It also may be of interest to calculate confidence intervals for the survival probability at a certain time after treatment. In this case, denote ψ as

$$\psi = \frac{\log t - x^T \beta}{b}$$

11

using the same patient as in the previous situation, at 500 days,

$$\hat{\psi} = \frac{\log 600 - (9.555 + 0.736 - 0.259 - 0.3052 - 0.064(5) - 1.695 - 0.823)}{1.370} = -0.359$$

Then using the fact that T follows a Weibull distribution, use ψ in the standard Weibull survivor function, $S(t) = e^{-e^x}$

$$\hat{S}(600) = e^{-e^{-0.3589301}} = 0.497369$$

This suggests that such a confidence interval will be centered around 0.497.

Since we wish to find a confidence interval for $S(600)$, it now remains to discern the variance of $\hat{\psi}$. This variance is given using the covariance matrix V provided by `survreg()`, adjusted so that it is in terms of β and b instead of the default β and $\log b$. This matrix will be a 12×12 matrix. Then the variance of ψ is given as

$$Var(\psi) = \frac{1}{\hat{b}^2}\left(\mathbf{x}, \frac{(log(600) - \hat{\beta}^T x)}{\hat{b}}\right) V \left(\mathbf{x}, \frac{(log(600) - \hat{\beta}^T x)}{\hat{b}}\right)^T = 0.0717$$

Making sure to the first term of $\mathbf{x}$ is 1 to account for the intercept.

Then

$$\frac{\hat{\psi} - \psi}{se(\hat{\psi})} \sim N(0, 1)$$

Meaning a 95% confidence interval for ψ is simply given as

$$\hat{\psi} \pm z_{0.95} se(\hat{\psi}) = (-0.884, -0.359)$$

meaning that a confidence interval for $S(600)$ is given as $(\exp(-0.884), \exp(-0.359)) = (0.413, 0.70)$

Inference and confidence intervals under the PH model are significantly more complicated to construct, and will thus not be done in this analysis.

Model Checking

It remains to check the fit of both of the proposed models, the adequacy of the AFT model is checked by comparing the Kaplan Meier estimates evaluated at the residuals, denoted $\hat{S}(\hat{\epsilon})$, and comparing this to the proposed distribution's survivor function evaluated at these same residuals, denoted $S_0(\hat{\epsilon})$. In this model, such a proposed distribution was a Weibull.

These residuals are caluclated using the fact that,

$$Y = \log T = x^T\beta + b\epsilon \implies \frac{\log T - x^T\hat{\beta}}{\hat{b}} = \hat{\epsilon}$$

Ideally, both curves should closely resemble each other

KM Estimate vs Weibull Estimate of Survival Function
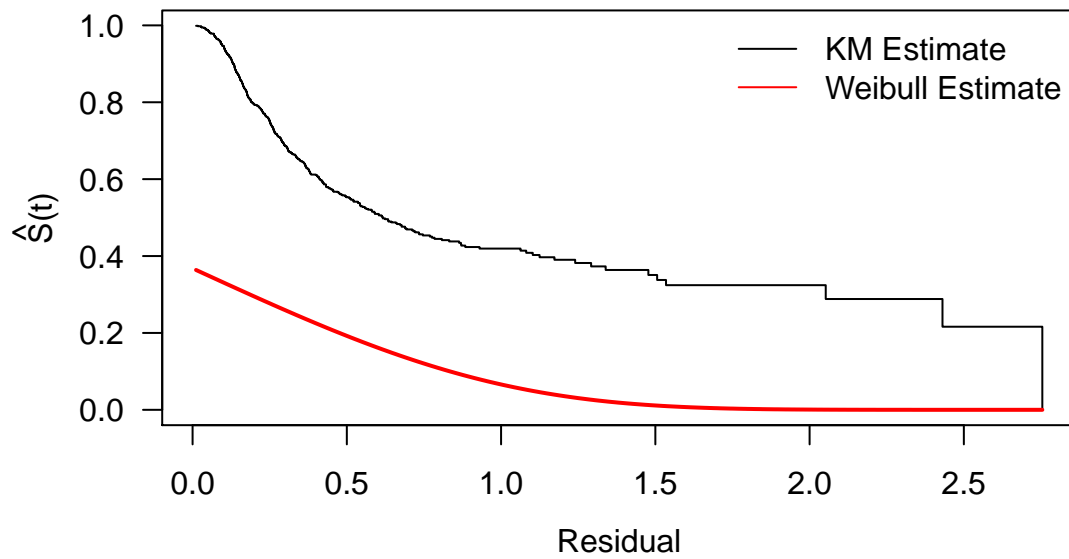
It is clear that these 2 distributions do not match each other in any significant way, we confirm this by plotting $\log(-log(\hat{S}(\hat{\epsilon}))$ versus the log-time associated with the KM fit. If the Weibull assumption was correct, this plot would be linear.

log(−log(S(t))) vs log(t)
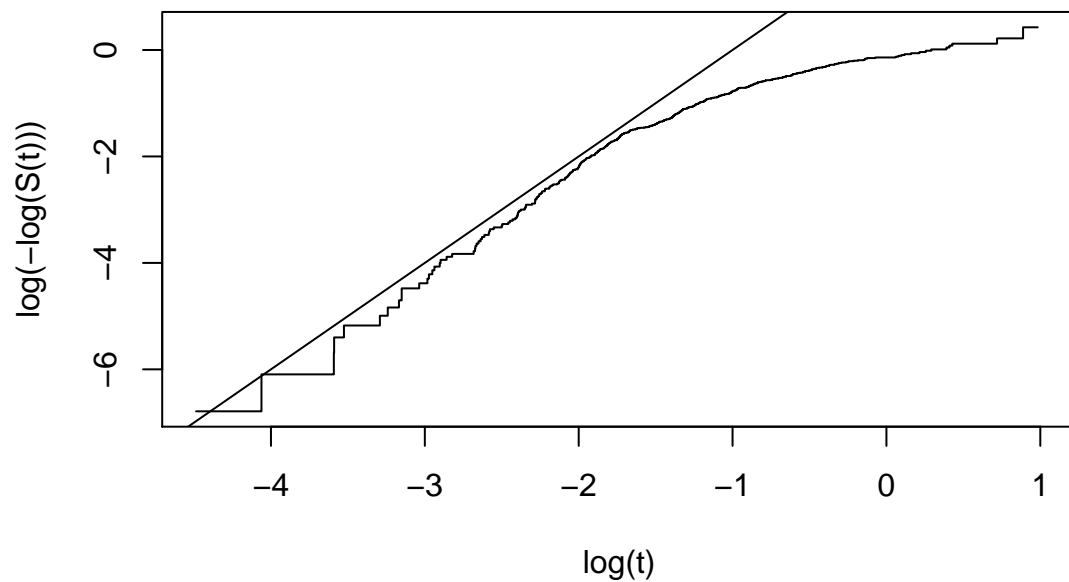
This linearity is clearly violated as time increases. This makes sense considering the first plot shown in this analysis. From that plot it was possible to deduce that as time increased to the scale of years, many

patient's cancer did not recur, indicating that they were in essence cured. This meant that since a traditional survivor function approaches 0 as time goes on (as was not the case here), a traditional model was unlikely to fit the data well.

In general it seems this AFT assumption is heavily flawed due to the fact that some patients may be cured. As such, it is not worthwhile to continue to investigate this model.

In order to check the adequacy of the Cox PH model, the primary assumption to check is that of proportional hazards. This is to say that hazards may be written as ratios of a baseline and a function of covariates, and that hazard ratio does not change over time, given that covariates remain the same. To test this, `cox.zph` was built to accept a `cox.ph` model and perform this test.

Table 11: Test of Proportionality

	rho	χ^2	p
rxLev	−0.020	0.183	0.669
rxLev+5FU	0.016	0.111	0.739
obstruct1	−0.121	6.750	0.0092 ***
adhere1	0.050	1.117	0.290
nodes	0.057	1.137	0.286
extent2	0.006	0.018	0.894
extent3	0.016	0.112	0.738
extent4	−0.008	0.029	0.865
surg1	0.063	1.759	0.185
node4	−0.152	9.548	0.002 ***
GLOBAL		22.104	0.015 **

From this test, it can be seen that `obstruct` and `node4` indicate non-proportionality. By examining the schoenfeld residuals associated with the variable `node4`, it can be seen that residuals tended to trend negatively

## Schoenfeld residuals over time of node4
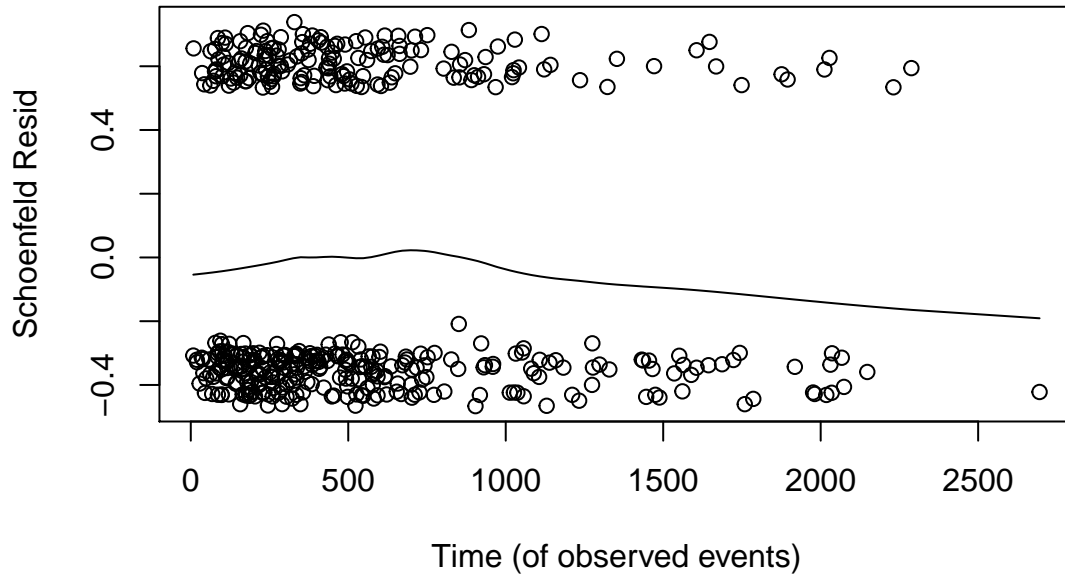
In terms of correcting this lack of proportionality, this suggests that `node4` and `obstruct` are time

varying. It is possible to fit another function of these covariates that depends on time, for example an indicator or a linear function of time, that attempts to correct for this. An indicator $I(t > 700)$ was added to the model to interact with `node4`, but this seemed to not offer a highly significant improvement. A linear function would likely improve upon this, but such a function and its implementation are beyond the scope of the course. This departure from proportionality may not be of serious concern, partially because these predictors are not of primary interest to this analysis.

With regard to the different types of residuals of this model: score, deviance, martingale, schoenfeld, and scaled schoenfeld, there sometimes existed patterns in the residuals, which were indicated by the inclusion of lowess lines.

In particular, the martingale residuals associated with `rx` appeared not to follow any pattern in general.
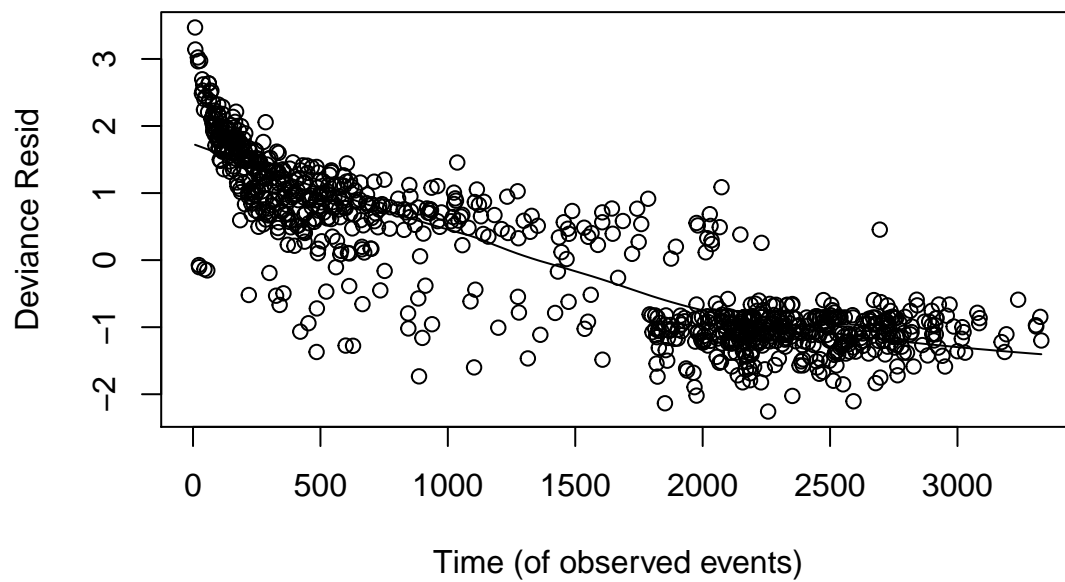
Martingale residuals over time of rx
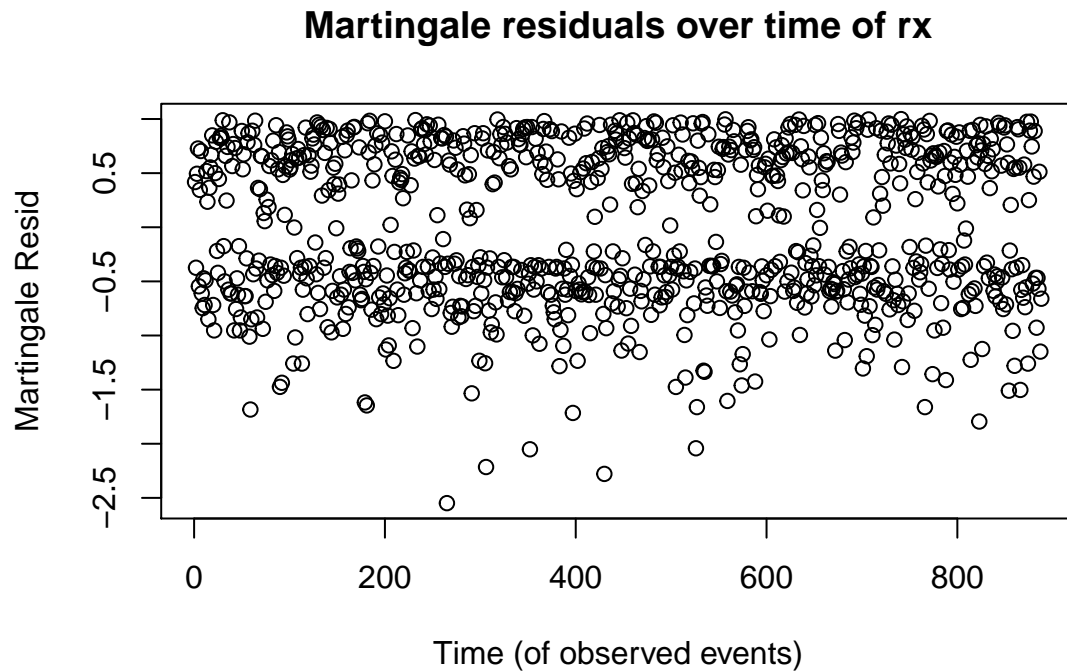
A further sample of the resdiual plots is shown below

Score residuals over time of rx
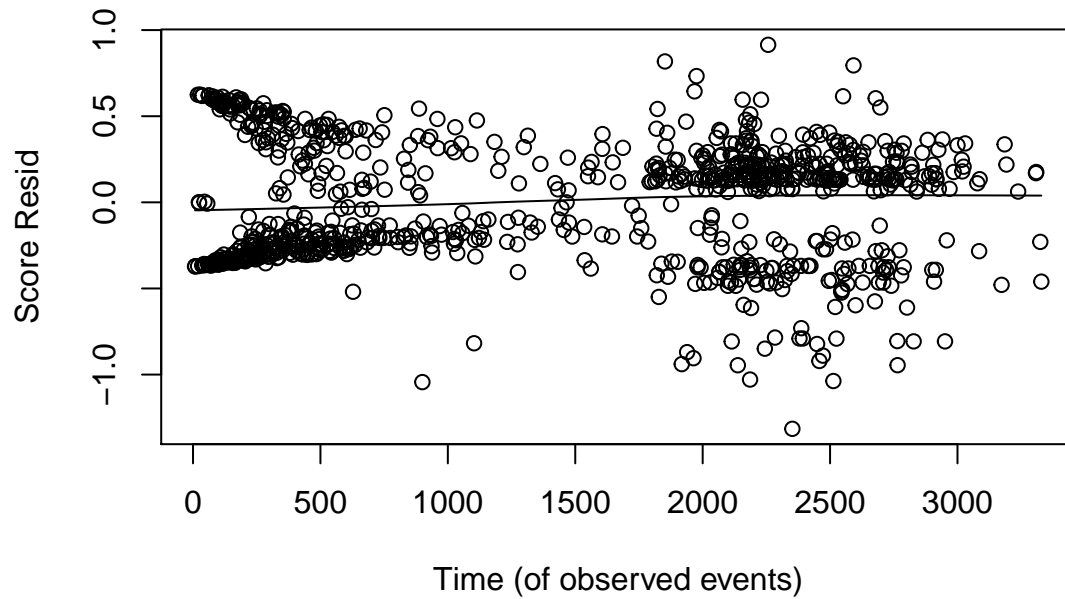

Deviance residuals over time of rx

The most important characteristic that should be checked in this model is its ability to fit the data. This is done by comparing the Kaplan Meier estimate of the survivor function to the Cox PH estimate of the survivor function.

```
##                    rho    chisq        p
## rxLev         -0.01213  0.06610 0.797097
```

```
## rxLev+5FU     -0.00344  0.00544 0.941225
## obstruct1     -0.11149  5.66611 0.017296
## adhere1        0.05418  1.32296 0.250061
## nodes          0.04913  1.07213 0.300465
## extent2        0.00193  0.00166 0.967519
## extent3        0.01338  0.08048 0.776643
## extent4       -0.00347  0.00543 0.941251
## surg1          0.06338  1.81557 0.177842
## node4          0.05183  1.43205 0.231430
## node4:ind.5001 0.14396  8.52586 0.003501
## GLOBAL              NA 37.04424 0.000113
```

```
plot(survfit(Surv(time,status)~1,data=data)$time, survfit(Surv(time,status)~1,data=data)$surv,type="s",
lines(survfit(model.PH3,type="aalen"), col="red")
legend("topright", c("KM Estimate", "Cox PH Estimate"),
    lty=1,col=c("black","red"), bty = "n")
```

KM Estimate vs Cox PH Estimate of Survival Function
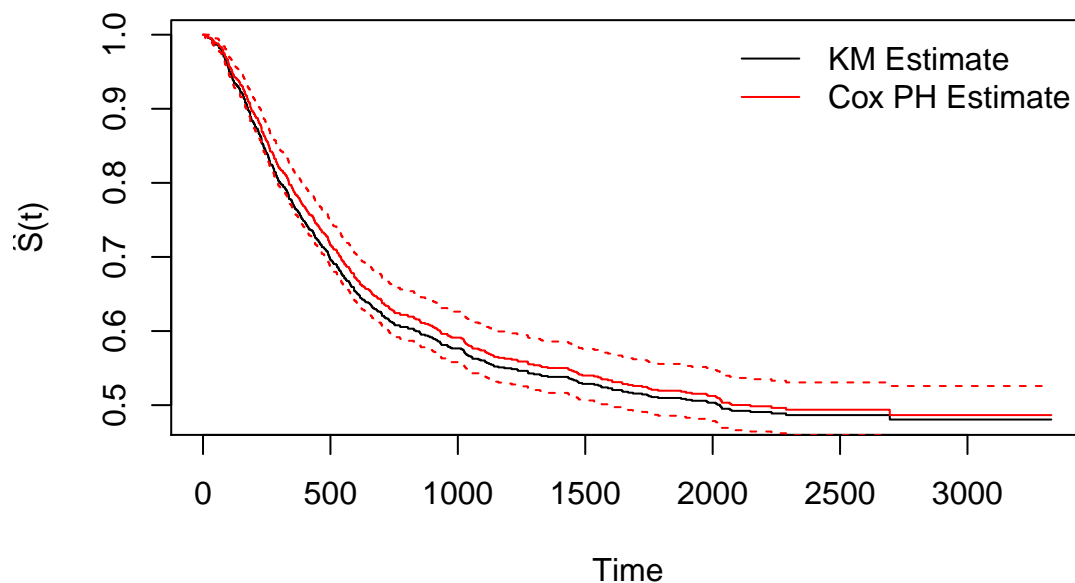
As can be seen, this model is able to emulate the Kaplan Meier estimate of the survivor function very well. It can be seen that the KM estimate always falls within the confidence bands assicated with the Cox PH estimate.

Mixture Model for a More Simple Problem

In lecture 4, the concept of a mixture model was discussed. In particular, the special case of a mixture model with 2 groups, $(1-p)\%$ of which have lifetimes much longer than the $p\%$ remaining group. This may apply in this case, since Kaplan Meier estimates of the survivor functions appear to level-off much before reaching 0. This would mean that these patients are likely to either never have a recurrence of cancer, or that such a recurrence would occur far in the future. Such a group, denoted group 2, would thus have a survivor function roughly equal to 1 on the time scale of group 1's recurrence times

$$S(t) = pS_1(t) + (1-p)S_2(t) = pS_1(t) + (1-p) \cdot 1 = pS_1(t) + (1-p) \qquad (*)$$

In the case of this dataset, we wish to have $S_1(t)$ be derived from an AFT model. Since different covariates may have different effects on this value of p, this model would be of the form

$$S(t|x) = p(x; \alpha)S_0(t|x; \beta) + (1 - p(x; \alpha)) \qquad (**)$$

These types of mixtures are difficult to attain through functions in R, however a package `flexsurvcure` by Jordan Amdahl was discovered that claims to be able to describe models[1]. This model was fit below,

AFT Cure Model vs KM Estimated Model
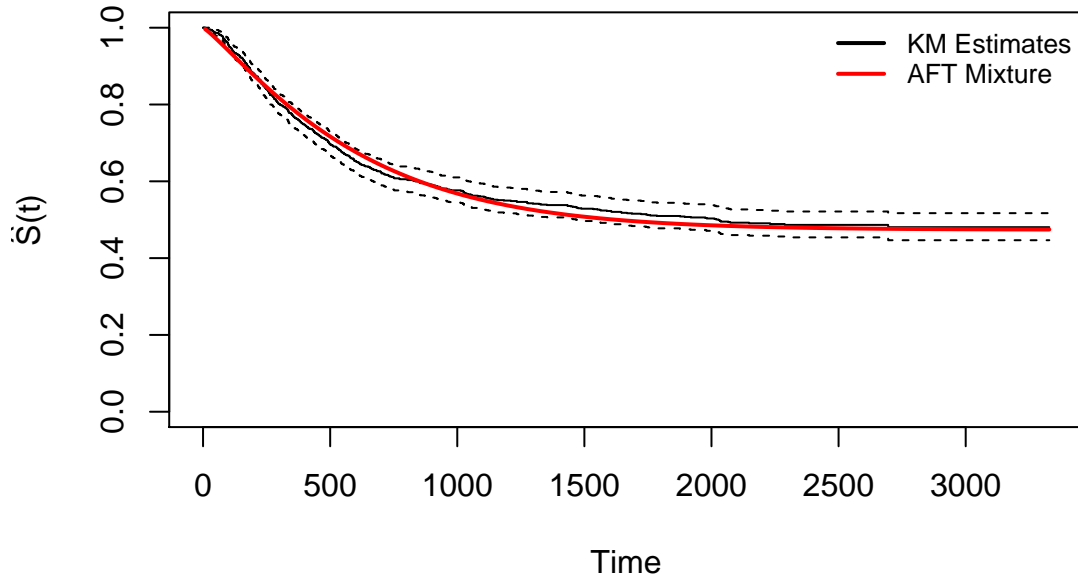
This model appears to have a near perfect fit to the data, however without knowing more about the methods applied in this function, caution should be used in using this function. Evidence to support such a fit is that Lawless investigates a similar situation in section 4.4 (without using any covariates, however) and achieves a similar result[4].

Given that this model is not supported natively in the `survival` package, and that this model is beyond the scope of this course, the more simple model $(*)$ was investigated, where we only wish to model the effect of levamisole and fluorouracil and observation seperately, in the absence of any covariates. In other words, 2 distinct Weibull(u_i, b_i) models are created, where u_1, b_1 are estimated using `rx=Lev+5FU` recurrence times and no covariates and u_2, b_2 are estimated using `rx=Obs` recurrence times with no covariates.

$$S_{Lev+5FU}(t) = pS_{Lev+5FU}(t) + (1-p)$$
$$S_{Obs}(t) = pS_{Obs}(t) + (1-p)$$

Since both of these models are of the form $(*)$, it also is necessary to estimate p in each case. The density function $f(t)$ for each model is given as the negative derivative with respect to t of $S(t)$

$$f(t) = -pS'(t)$$

.

The estimates of u_i, b_i, p were found by maximizing the log likelihood function associated with this model

$$\prod_i \left(pS(t_i) + (1-p)\right)^{1-\delta_i} \left(f(t_i)\right)^{\delta_i}$$

```r
#Making 2 new data sets, with only 1 type of rx each
data.simple.fu <- data[data$rx=="Lev+5FU",]
data.simple.obs <- data[data$rx=="Obs",]

data.simple.fu$rx <- factor(data.simple.fu$rx)
data.simple.obs$rx <- factor(data.simple.obs$rx)

#vs 1 data set with 2 types of rx, for a simple AFT model
data.simple.reg <- data[data$rx!="Lev",]
data.simple.reg$rx <- factor(data.simple.reg$rx)
model.simple <- survreg(Surv(time,status)~rx,data=data.simple.reg)

km.estimates <- survfit(Surv(time,status)~rx,data=data.simple.reg)


#mixture log likelihood
Mix.Lik.p<-function(t,param,status){
  p=param[1]
  u=param[2]
  b=param[3]
-sum(log((p/b*exp((log(t)-u)/b)*exp(-exp((log(t)-u)/b)))^(status)*(p*exp(-exp((log(t)-u)/b))+(1-p))^(1-status)))
}

param.est.fu=nlminb(start=c(0.5,5,0.5),
                    Mix.Lik.p,
                    status=data.simple.fu$status,
                    t=data.simple.fu$time)$par

param.est.obs=nlminb(start=c(0.5,5,0.5),
                     Mix.Lik.p,
                     status=data.simple.obs$status,
                     t=data.simple.obs$time)$par

#survival probabilities using MLE
st=param.est.fu[1]*exp(-exp((log(data.simple.fu$time)-param.est.fu[2])/param.est.fu[3]))+(1-param.est.fu[1])
st2=param.est.obs[1]*exp(-exp((log(data.simple.obs$time)-param.est.obs[2])/param.est.obs[3]))+(1-param.est.obs[1])

df.test <- data.frame(time=data.simple.fu$time, surv=st)
df.test <- df.test[order(df.test$time),,drop=F]

df.test2 <- data.frame(time=data.simple.obs$time, surv=st2)
df.test2 <- df.test2[order(df.test2$time),,drop=F]

par(las=1)
plot(km.estimates, main="Mixture Model vs. Traditional Weibull AFT",
     ylab =c(expression(paste(hat(S),"(t)"))),xlab="Time")
lines(predict(model.simple,
              newdata=list(rx="Obs"),
              type="quantile",
              p=seq(0,1,by=.01)),
      seq(1,0,by=-.01),
      col="darkgreen")

lines(predict(model.simple,
              newdata=list(rx="Lev+5FU"),
              type="quantile",
              p=seq(0,1,by=.01)),
      seq(1,0,by=-.01),
      col="darkgreen")
lines(df.test$time,df.test$surv,type="l",col="red")
lines(df.test2$time,df.test2$surv,type="l",col="red")
legend("topright", legend=c("KM Estimates", "AFT", "Mixture"),
       col=c("black", "darkgreen", "red"), lty=1, cex=0.8,lwd=2,bty="n")
```
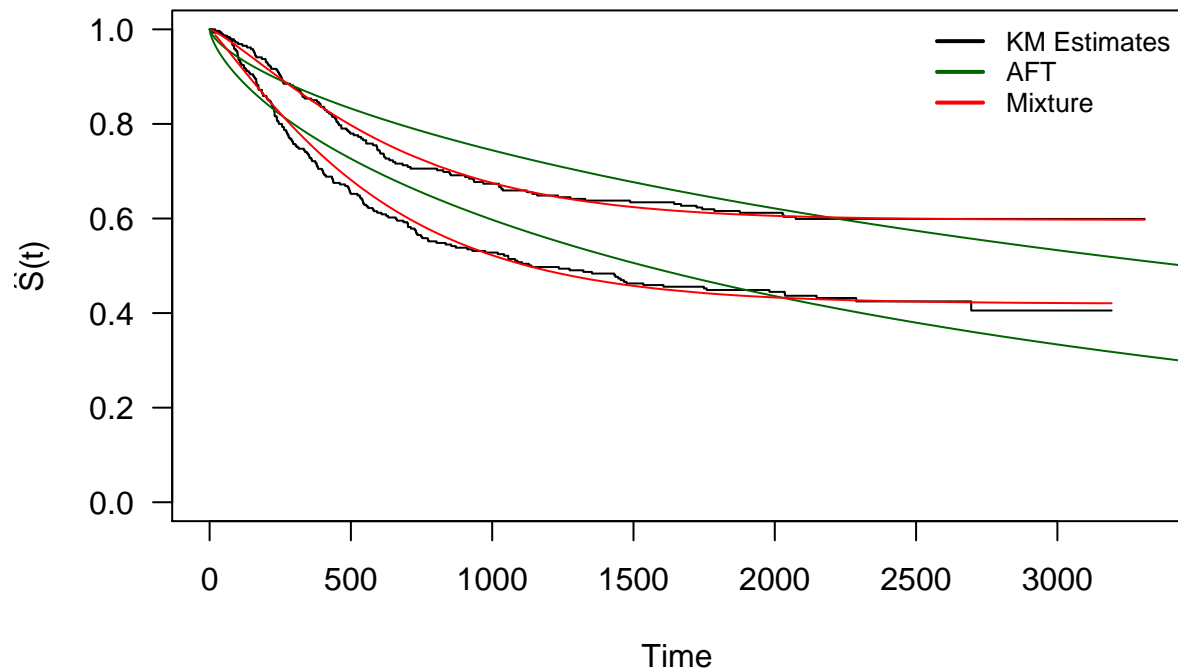
Mixture Model vs. Traditional Weibull AFT

This could be taken further to calculate the respective variances of these estimates, and these could further be used to make inference upon either those observed or those under the levamisole and fluorouracil therapy. This kind of analysis lies beyond the content of this course, however.

Conclusion

To conclude, the ordinary AFT assumption of this model was not well founded, and this model has a very poor fit to the data. This suggests that value it could provide is minimal, if it can provide any at all. However, there does seem to be potential in using a mixture AFT model where a certain percentage of patients can be classified as cured, and will never have their cancer reoccur.

The semi-parametric Cox PH model seems to be able to fit the data much better, although there may exist issues in assuming the proportionality of the hazard function under some of the covariates. Unlike the AFT model, however, this PH model is much more difficult to make inference upon. This model does, however, agree with the results obtained by the original analysis done by Laurie et al.

In general, however, it seems safe to conclude that the most important covariate of interest, `rx` did have differing levels, primarily `rx=Lev+5FU`, where recurrence times were increased significantly under this therapy when compared to no therapy, or merely the administering of just levamisole. `rx=Lev` itself appeared to have no statistically significant effect when compared to mere observation, which was of note.

Other covariates that were significant in affecting recurrence times were `surg`, `nodes` (and `node4`), and `extent`. These covariates all suggested that as the disease becomes more involved, and treatment is delayed, the seriousness of the disease is increased, meaning that the time to recurrence of the cancer is decreased.

Recommendations

In the future, there are 2 distinct paths that should be of interest in future analysis. The first is attempting to create an AFT mixture model with all significant covariates, as opposed to the simple attempt

with a single level of a covariate that was attempted in this analysis. This model would be powerful in that it may be easier to make inference and interpret than the alternative Cox PH model.

The second path of interest should be to finely tune the current Cox PH model to attempt to correct for the 2 covariates that appear to violate the assumption of proportionality. Instead of this, it would also be possible to stratify by these variables if they were deemed important enough to continue to consider.

Limitations

Further inspection of the residuals associated with the Cox PH model, as well as an inspection as to the seriousness of the lack of proportionality would impose limits on the ability to make inference using this model.

References

[1] Amdahl J. flexsurvcure: Flexible Parametric Cure Models. CRAN ˘ Package flexsurvcure. 2019. Cran.r: https://cran.r-project.org/web/packages/flexsurvcure/?fbclid=IwAR3f0T-WaOPnDo-2WMZBFNrMIhja6hWSQ6HOQKxQpBSfksNiKesfQpkr9rU

[2] Breheny P. Accelerated Failure Time Models: Survival Data Analysis (BIOS 7120). The University of Iowa. $2019; 1 : 25$. myweb: hhttp://myweb.uiowa.edu/pbreheny/7210/f15/notes/10-15.pdf?fbclid=IwAR0ugtQvHLUYWUIsuxZSHIkm_gvO5tChcMo3gfIeDBy9YfNtS8bVoWvn150

[3] JA Laurie, CG Moertel, TR Fleming, HS Wieand, JE Leigh, J Rubin, GW McCormack, JB Gerstner, JE Krook and J Malliard. Surgical adjuvant therapy of large-bowel carcinoma: An evaluation of levamisole and the combination of levamisole and fluorouracil: The North Central Cancer Treatment Group and the Mayo Clinic. J Clinical Oncology, $7 : 1447 − 1456, 1989$.

[4] Lawless JF. Statistical Models and Methods for Lifetime Data: 2nd Edition. John Wiley & Sons, $4 : 181 − 185, 2003$.

[5] Lawless JF. Statistical Models and Methods for Lifetime Data: 2nd Edition. John Wiley & Sons, $6 : 273, 2003$.

[6] Moertel CG, Fleming TR, Macdonald JS, et al. Fluorouracil plus Levamisole as Effective Adjuvant Therapy after Resection of Stage III Colon Carcinoma: A Final Report. Ann Intern Med. $1995; 122 : 321 ˘ 326. doi :$ hhttps://doi-org.proxy1.lib.uwo.ca/10.7326/0003-4819-122-5-199503010-00001

Appendix

Roles

Daniel Grant: Model Checking, Analysis, Mixture Model, Overall Editing
Ruobing Wei: Introduction, Preliminary Plots, Initial Modeling
Sana Mungroo: References, Conclusion, Overall Editing

A: Data

```r
library(survival)
library(tidyverse)
library(e1071)
library(GGally)
library(ggfortify)
library(ggplot2)
library(flexsurv)
library(kableExtra)
```

```r
library(stargazer)
library(cowplot)
library(flexsurvcure)
data <- colon[seq(2,1858,2),]
data <- na.omit(data)
data$sex     <- factor(data$sex)
data$obstruct <- factor(data$obstruct)
data$perfor  <- factor(data$perfor)
data$adhere  <- factor(data$adhere)
data$differ  <- factor(data$differ)
data$extent  <- factor(data$extent)
data$surg    <- factor(data$surg)
```

B: Code

```r
#NP model of rx
nonpara.model <- survfit(Surv(time,status)~rx, data = data, type = "kaplan-meier")
par(mgp=c(2.5,1,0))
plot(nonpara.model,
     main="Kaplan Meier Survival Probability Estimates",
     ylab =c(expression(paste(hat(S),"(t)"))),
     xlab="Time",
     col=c("black","red","blue"))
legend("topright",
       legend=c("Obs", "Lev", "Lev+5FU"),
       col=c("black", "red", "blue"),
       lty=1,
       cex=0.8,
       lwd=2,
       bty="n")


#survdiff on rx
survdiff(Surv(time,status)~rx, data = data)



#AFT MODEL
model.AFT <-survreg(Surv(time,status)~rx + sex + age + obstruct + perfor + adhere +
                                    nodes + differ + extent + surg + node4,
                 data = data,
                 dist = "weibull")
summary(model.AFT)

#Weibull Para PH model
model.PH=flexsurvreg(Surv(time,status)~rx + sex + age + obstruct + perfor + adhere +
                                    nodes + differ + extent + surg + node4,
                 data = data,
                 dist = "weibullPH")
summary(model.PH)

#Reduced AFT
model.AFT2 <-survreg(Surv(time,status)~ rx + obstruct +  adhere +
                                    nodes + extent + surg + node4,
                 data = data,
                 dist = "weibull")
summary(model.AFT2)
```

```r
#Full PH
model.PH2 <-coxph(Surv(time,status)~rx + sex + age + obstruct + perfor + adhere +
                                 nodes + differ + extent + surg + node4,
                    data = data)
summary(model.PH2)

#reduced model
model.PH3 <-coxph(Surv(time,status)~rx + obstruct + adhere +
                                 nodes + extent + surg + node4,
                    data = data)

summary(model.PH3)

#Reduced AFT covariarnace matrix
cv.mx <- model.AFT2$var
cv.mx[,12]<-cv.mx[,12]*model.AFT2$scale
cv.mx[12,]<-cv.mx[12,]*model.AFT2$scale

psi.hat <- (log(600)-(9.555+0.736-0.259-0.3052-0.064*(5)-1.695-0.823))/model.AFT2$scale
var.psi <- 1/model.AFT2$scale^2*c(1,0,1,1,1,5,0,0,1,0,1,psi.hat)
           %*%cv.mx%*%
           t(t(c(1,0,1,1,1,5,0,0,1,0,1,psi.hat)))

psi.hat+qnorm(0.975)*sqrt(0.07171199)
psi.hat-qnorm(0.975)*sqrt(0.07171199)


#Model Checking AFT
model.AFT2.res<-exp(-model.AFT2$linear.predictor/model.AFT2$scale)* (Surv(data$time, data$status)[,1])^(1/model.AFT
np.fit.res<-survfit(Surv(model.AFT2.res,data$status)~1)
par(mgp=c(2.5,1,0))
par(las=1)
#par(mar=c(5,6,4,2)-1)
plot(np.fit.res$time,np.fit.res$surv,
     type="s",
     xlab="Residual",
     ylab=c(expression(paste(hat(S),"(t)"))),
     main="KM Estimate vs Weibull Estimate of Survival Function")
x <- seq(min(model.AFT2.res), max(model.AFT2.res), length.out = 400) #adding a smooth plot of S(ehat)
y <- exp(- exp(x))
lines(x, y, col = "red", lwd = 2)
legend("topright",
       c("KM Estimate", "Weibull Estimate"),
       lty=1,
       col=c("black","red"),
       bty = "n")

#Mode AFT model checking
np.fit.res<-survfit(Surv(model.AFT2.res,data$status)~1)
plot(log(np.fit.res$time),
     log(-log(np.fit.res$surv)),
     type="s",xlab="log(t)",
     ylab="log(-log(S(t)))",
     main = "log(-log(S(t))) vs log(t)")
abline(2,2)


#AFT residual plots (not in analysis)
```

```r
de.res<-residuals(model.AFT2, type="deviance")

plot(model.AFT2$linear.predictor,
     de.res,
     main = "Residuals vs Linear Predictors",
     xlab = "Linear Predictor",
     ylab="Residual")

plot(Surv(data$time, data$status)[,1],
     de.res,
     main = "Residuals vs Time",
     xlab = "Time (days)",
     ylab="Residual")


#PH model checking
cox.zph(model.PH3)
plot(data$time[data$status==1],
     residuals(model.PH3, "schoenfeld")[,10],
     main = "Schoenfeld residuals over time of node4",
     xlab = "Time (of observed events)",
     ylab = "Schoenfeld Resid")
lines(lowess(data$time[data$status==1],
             residuals(model.PH3, type="schoenfeld")[,10]))

#attempting to fix time varying node4
data$ind.500 <-factor(as.numeric(data$time>700))

model.PH4 <- coxph(Surv(time,status)~rx +obstruct + adhere +
                                      nodes + extent + surg + +node4:ind.500+node4,
                   data = data)

cox.zph(model.PH4)

#back to residual checking in reduced model
#no pattern mean 0.
plot(residuals(model.PH3, "martingale"),
     model.PH3$time,
     main = "Martingale residuals over time of rx",
     xlab="Time (of observed events)",
     ylab="Martingale Resid")

plot(data$time,
     residuals(model.PH3, "score")[,1],
     main = "Score residuals over time of rx",
     xlab="Time (of observed events)",
     ylab="Score Resid")
  lines(lowess(data$time, residuals(model.PH3, type="score")[,1]))

plot(data$time, residuals(model.PH3, "deviance"),
     main = "Deviance residuals over time of rx",
     xlab="Time (of observed events)",
     ylab="Deviance Resid")
lines(lowess(data$time, residuals(model.PH3, type="deviance")))

plot(futime[fustat==1],
     residuals(cox.fit, type="scaledsch"),
     xlab="Time",
```

```r
    ylab="Residual",
    main="scaledsch")
lines(lowess(futime[fustat==1], residuals(cox.fit, type="scaledsch")))

#KM vs COXPH
plot(survfit(Surv(time,status)~1,data=data)$time,
    survfit(Surv(time,status)~1,data=data)$surv,
    type="s",
    xlab="Time",
    ylab=c(expression(paste(hat(S),"(t)"))),
    main="KM Estimate vs Cox PH Estimate of Survival Function")

lines(survfit(model.PH3,type="aalen"), col="red")
legend("topright",
    c("KM Estimate", "Cox PH Estimate"),
    lty=1,
    col=c("black","red"),
    bty = "n")


#CURE MODEL
cure_model2 <- flexsurvcure(Surv(time,status)~rx + sex + age + obstruct + perfor + adhere +
                                nodes + differ + extent + surg + node4, data=data,
                        link="logistic",
                        dist="weibull",
                        mixture=T)
#cure_model2

plot(cure_model2,
    main = "AFT Cure Model vs KM Estimated Model",
    xlab="Time",
    ylab=c(expression(paste(hat(S),"(t)"))))
legend("topright",
    legend=c("KM Estimates", "AFT Mixture"),
    col=c("black", "red"),
    lty=1,
    cex=0.8,
    lwd=2,
    bty="n")

#Making 2 new data sets, with only 1 type of rx each
data.simple.fu <- data[data$rx=="Lev+5FU",]
data.simple.obs <- data[data$rx=="Obs",]

data.simple.fu$rx <- factor(data.simple.fu$rx)
data.simple.obs$rx <- factor(data.simple.obs$rx)

#vs 1 data set with 2 types of rx, for a simple AFT model
data.simple.reg <- data[data$rx!="Lev",]
data.simple.reg$rx <- factor(data.simple.reg$rx)
model.simple <- survreg(Surv(time,status)~rx,data=data.simple.reg)

km.estimates <- survfit(Surv(time,status)~rx,data=data.simple.reg)

#SIMPLE MIXTURE MODEL
#mixture log likelihood
Mix.Lik.p<-function(t,param,status){
  p=param[1]
```

```r
  u=param[2]
  b=param[3]
-sum(log((p/b*exp((log(t)-u)/b)*exp(-exp((log(t)-u)/b)))^(status)*(p*exp(-exp((log(t)-u)/b))+(1-p))^(1-status)))
}

param.est.fu=nlminb(start=c(0.5,5,0.5),
                    Mix.Lik.p,
                    status=data.simple.fu$status,
                    t=data.simple.fu$time)$par

param.est.obs=nlminb(start=c(0.5,5,0.5),
                     Mix.Lik.p,
                     status=data.simple.obs$status,
                     t=data.simple.obs$time)$par

#survival probabilities using MLE
st=param.est.fu[1]*exp(-exp((log(data.simple.fu$time)-param.est.fu[2])/param.est.fu[3]))+(1-param.est.fu[1])
st2=param.est.obs[1]*exp(-exp((log(data.simple.obs$time)-param.est.obs[2])/param.est.obs[3]))+(1-param.est.obs[1])

df.test <- data.frame(time=data.simple.fu$time, surv=st)
df.test <- df.test[order(df.test$time),,drop=F]

df.test2 <- data.frame(time=data.simple.obs$time, surv=st2)
df.test2 <- df.test2[order(df.test2$time),,drop=F]

par(las=1)
plot(km.estimates, main="Mixture Model vs. Traditional Weibull AFT",
        ylab =c(expression(paste(hat(S),"(t)"))),xlab="Time")
lines(predict(model.simple,
              newdata=list(rx="Obs"),
              type="quantile",
              p=seq(0,1,by=.01)),
      seq(1,0,by=-.01),
      col="darkgreen")

lines(predict(model.simple,
              newdata=list(rx="Lev+5FU"),
              type="quantile",
              p=seq(0,1,by=.01)),
      seq(1,0,by=-.01),
      col="darkgreen")
lines(df.test$time,df.test$surv,type="l",col="red")
lines(df.test2$time,df.test2$surv,type="l",col="red")
legend("topright", legend=c("KM Estimates", "AFT", "Mixture"),
        col=c("black", "darkgreen", "red"), lty=1, cex=0.8,lwd=2,bty="n")



######ADDITIONAL CODE TO PLOT KM ESTIMATES FOR EACH SINGLE COVARIATE MODEL
plot1<-autoplot(test.model11,
         conf.int = FALSE,
         censor.size = 2,
         surv.size = 1.2,
         ylab = c(expression(paste(hat(S),"(t)"))),
         xlab = "Time (days)",
         main = "KM",
         legend=F)+
  labs(color="Trt") +
```

```r
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
plot3<-autoplot(test.model3,
        conf.int = FALSE,
        censor.size = 2,
        surv.size = 1.2,
        ylab = "",
        xlab = "Time (days)",
        main = "KM",
        legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))

plot4<-autoplot(test.model4,
        conf.int = FALSE,
        censor.size = 2,
        surv.size = 1.2,
        ylab = "",
        xlab = "Time (days)",
        main = "KM",
        legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
plot5<-autoplot(test.model5,
        conf.int = FALSE,
        censor.size = 2,
        surv.size = 1.2,
        ylab = c(expression(paste(hat(S),"(t)"))),
        xlab = "Time (days)",
        main = "KM",
        legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
plot7<-autoplot(test.model7,
        conf.int = FALSE,
        censor.size = 2,
        surv.size = 1.2,
        ylab = "",
        xlab = "Time (days)",
        main = "KM",
        legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
plot8<-autoplot(test.model8,
        conf.int = FALSE,
        censor.size = 2,
        surv.size = 1.2,
        ylab = "",
        xlab = "Time (days)",
        main = "KM",
        legend=0)+
  labs(color="Trt") +
  theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
```

```r
plot9<-autoplot(test.model9,
         conf.int = FALSE,
         censor.size = 2,
         surv.size = 1.2,
         ylab = c(expression(paste(hat(S),"(t)"))),
         xlab = "Time (days)",
         main = "KM",
         legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))
plot10<-autoplot(test.model10,
         conf.int = FALSE,
         censor.size = 2,
         surv.size = 1.2,
         ylab = c(expression(paste(hat(S),"(t)"))),
         xlab = "Time (days)",
         main = "KM",
         legend=F)+
  labs(color="Trt") +
    theme(legend.key.size = unit(0.5, "cm"))+
  theme(axis.title.y = element_text(angle=0,vjust=0.5))

plot_grid(plot1,plot3,plot4,plot5,plot7,plot8,plot9,plot10, ncol=3,nrow=3)
```