Survival Analysis Final Report: Data Analysis for the Sequential Primary Biliary Cirrhosis Data

Ya-fang Yan

Abstract

This article presents an application of the Kaplan-Meier estimator and a real data, the sequential promary biliary cirrhosis collected in Mayi clinic, which contains several time dependent covariates and the observations is measured repeatedly fits for the Cox proportional hazard models because the Cox PH model is the most popular method for survival data. Also, by employing the model fit criterion to determine a suitable model for the real data and the criterion considered were the Akaike information criterion, called AIC for short. We find some factors, which have time to event, are agreed the assumptions of PH model. So the Cox PH model is appropriate to the data and the lifetime of patients would be related strongly to the covariates bilirubin, age, presence of edema, SGOT, presence of hepatomegaly, prothrombin time, log(platelets), albumin and log(alkaline). Moreover, whether patients took D-penicillamine or not would not affect clealy the lifetime of patients with PBC.

Key Words: Cox proportional hazard model; Kaplan-Meier estimate; AIC; Time-dependent.

1. Introduction

We consider methods for the analysis of data when the response of interest is the time until some event occurs, such events are generically referred to as failure. A principal problem examined is that of developing methods for assessing the dependence of failure time on explanatory variables. Most of the statistical research was concentrated on parametric models. While the survival analysis attempts to cover both the parametric and nonparametric methods, the emphasis is on the more recent nonparametric developments with applications to medical research.

This data set is a followup to the original PBC data set, which contains the baseline measurements and survival of 426 subjects, 312 formal study participants, and 106 eligible nonenrolled subjects. This data set contains multiple laboratory results, but only on the first 312 patients, among which 140 had died and the rest were censored and the sex ratio is at least $9:1$ (women to men) as of the data set. The main purpose of this study is to investigate the impact of D-penicillamine and bilirubin

to lifetime of patients with Primary Biliary Cirrhosis (PBC). The data set contains the covariates which are drug, patient's age at registration, patient's sex, presence of ascites, hepatomegaly, spiders and edema, serum bilirubin, albumin, alkaline phosphatase, serum glutamic-oxaloacetic transaminase (SGOT), platelets per cubic, prothrombin time and histologic stage of disease, etc...

Chapter 2 gives a summarized account of the techniques for analysis, such as the Kaplan-Meier estimator, AIC and Cox's PH model for time dependent covariate. Checking the PH assumption and applying the methodologies in chapter 2 to the PBC data are covered in chapter 3. In chapter 4, there are explanations of the suitable Cox's PH model obtained by the criteria AIC. Finally, we make some conclusions about the model and some extensions are suggested by chapter 5.

2. The Proposed Estimators

2.1. Kaplan-Meier Estimator

The standard nonparametric technique to estimate the survival function is proposed by Kaplan and Meier (1958), is called the Product-limit estimator. Let $t_1 < t_2 < \cdots < t_D$ represent the observed failure times in a sample of size n and Y_i be the number of individuals at risk prior to time t_i. This estimator is defined as

$$\hat{S}(t) = \prod_{t_i \leq t} \left[1 - \frac{d_i}{Y_i}\right]$$

for all t in the range where d_i means the number of events at time t_i. The Product-limit estimator is a right continuous step function with jumps at the observed event times and it provides an efficient means of estimating the survival function for right-censored data. An alternative nonparametric estimator is suggested by Nelson (1972). It has better small-sample-size performance estimator based on the Product-limit estimator.

On the other hand, we wish to test whether the survival functions of two or more samples could have significant difference which is called log rank test. at risk in combined sample at time t_i. Let O_i be the observed numbers and E_i be the expected numbers of failures in group i, $i = 1, \cdots, K$. Then the log rank test statistic is defined as

$$\mathcal{X}^2 = \sum_{i=1}^{K} \frac{(O_i - E_i)^2}{E_i}$$

which has approximately Chi-square distribution with degree of freedom $K - 1$. A large value $\mathcal{X}^2$ could lead to reject the testing hypothesis that there are discrepancies in survivor among the K families.

In the future, we use the Product-limit estimator to make crude comparisons the survival curves with time independent covariates in PBC data. We can find if there are differences in survival among the covariate's classifications by plotting survival curves and computing the $\mathcal{X}^2$ obtained from log-rank test. By the next step, we use semi-parametric technique to check these conditions, and hence, Cox's PH model is applied.

2.2. Cox's Proportional Hazard Regression model

The Cox's regression model (Cox, 1972) is widely used in epidemiological research to examine the association between an exposure and a health outcome. In a typical approach to the analysis of epidemiologic data with a continuous exposure variable, the exposure is transformed to an ordinal or nominal polytomous variable and relative risk (RR) is modeled as a step function of the exposure. The Cox's model is used to analyze censored data. Suppose the observed data are the triples (t_i, z_i, c_i) where t_i is the possibly censored survival time, z_i the scalar predictor variable, and c_i the event indicator, taking values of 1 if the event of interest occurred and 0 if it did not. Then, the Cox's proportional hazard model takes the form

$$h(t|\mathbf{Z}) = h_0(t) \exp\{Z'\beta\}, \tag{1}$$

where $h_0(.)$ is an arbitrary unspecified hazard function and β is the regression coefficient. If $h_0(t) = \lambda$ then (1) reduces to the exponential regression model; the Weibull model is the special case $h_0(t) = \lambda\gamma(\lambda\gamma)^{\gamma-1}$.

In other instances, the covariate $\mathbf{Z}(t)$ may be thought of as a stress factor affecting the individuals under study at time t. With such time-dependent covariates, the Cox's proportional hazard model is of the form

$$h(t|\mathbf{Z}(t)) = h_0(t) \exp\{Z(t)'\beta\}. \tag{2}$$

In next part, we would like to apply the PBC data to Cox's proportional hazard model and check whether the assumptions of proportional hazard holds. There are some methods proposed for diagnosing and assessing the proportional hazard assumptions. First, there has a program provided in R, called **cox.zph**, which supplies a method to test the assumption of proportional hazards. Furthermore, the object **cox.zph** contains an output of the martingale residuals. If the assumption is true then the regression is constant over time and the plot of time versus martingale residuals would be a horizontal line. Second, the Cox-Snell residuals is used to determine whether the fitted model is suitable for data or not. If the plot of residuals versus the estimated cumulative hazard should be a straight line thought the original with the slope 1 then the Cox's model fits the data. Finally, the deviance residuals can be used for identifying poorly predicted subjects. We can obtain

there exist other patterns which are not added to the Cox's model from the plot of individuals against the deviance residuals.

2.3. Akaike Information Criterion

The most popular and readily available methods for model fit criteria are the Akaike information criterion, AIC for short. The AIC method estimates the expected Kullback-Leibler (KL) information (Kullback and Leibler, 1951), a measure of the information lost when using an approximating distribution for estimation and inference instead of the true (unknown) distribution. The degree of freedom of the model give a bias correction to the expected KL information in large samples and act as a penalty on the numbers of parameters in the model. The optimal model minimizes AIC with respect to degree of freedom providing a balance between model fit (via the log-likelihood) and parsimony (df).

A popular alternative model selection criterion is the Bayesian information criterion or BIC. The BIC estimates the Bayesian factor comparing candidate models to one another and can be applied even when no prior distribution are explicitly specified. The definitions of AIC and BIC are as follow :

Parameter selection in terms of model degree of freedom

Criterion	Fromular
AIC	$-2\log L + 2df$
BIC	$-2\log L + df\log n$

df=model effective degrees of freedom;
n=total sample size;
$\log L$=log partial likelihood.

Although the equations of AIC and BIC look very similar, they originate from quite different frameworks. The BIC assumes that the true generation model is in the set of candidate models and it measures the degree of brief that a certain model is the true data-generating model. The AIC dose not assume that any of the candidate models is necessarily true. And these has code, **stepAIC**, included in the R and we can set k is equal to 2 if the AIC and $\log(n)$ if the BIC for the function in R codes. However, we use the AIC method to selection a fit model for the PBC data because the AIC is much convenient to compute than the BIC. Hence, we will analysis the PBC data with applying the AIC to choose a suitable Cox's model.

3. The data analysis for sequential PBC

We would not talk about the problem of missing data in this present and the covariate cholesterol contains much more missing data, so we intend to ignore this covariate. And we would make log-transformation for covariates alkaline and platelets

before analyzing this data. Since time independent covariates in sequential PBC data are only drug and sex, we could compute the Kaplan-Meier estimator to estimate the survival functions for these two covariates.

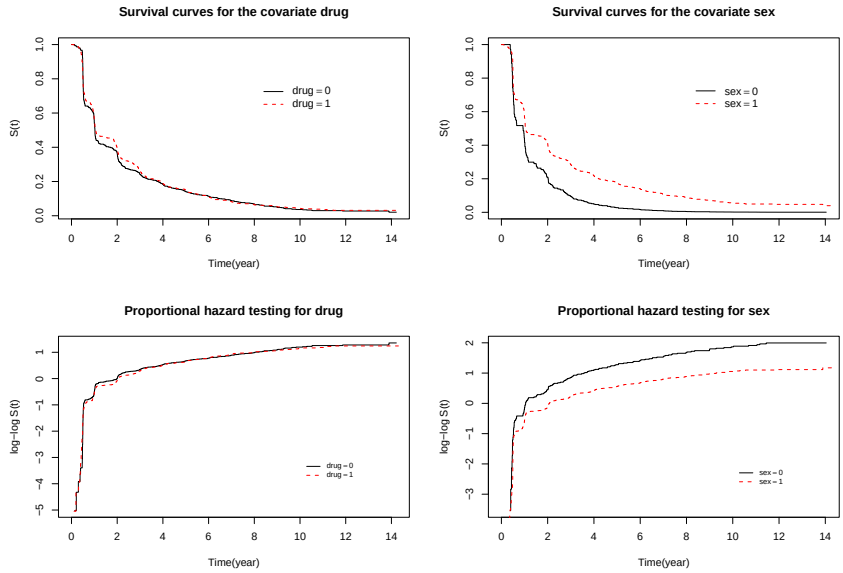

Figure 1: The survival and hazard functions for the patients of time independent covaraites.

Figure 1 plots the estimators of survival function for each level of drug and sex with the Kaplan-Meier estimator. In figure 1 we know that there are discrepancies for the patients between male and female and the female patients have more larger survival probability than male. Contrary to drug, the survival functions of the two drug levels look like similar to each other classification. There are a little cross over within each level of the two covaraites, but we also intend to say that the cross over dose not exist in this data. And the figures of time versus log-log survival probabilities for covariates drug and sex in figure 1 are shown the same results which the assumptions of proportional hazards is true for covariate sex. Therefore, the covariate drug seems to be the same curv for each drug level and the cross over dose not exist within each levels of drug and sex. On the other hand, we can compute the values of log rank test for drug and sex by using the code **survdiff** in R program and their p-values obtained by log rank test are 0.992 (larger than 0.05) and 0.00243 (less than 0.05), respectively. So the covariate drug might not have differences between two drug levels but there are significantly discrepancies for sex. Nevertheless, the rest covariates are time dependent and the AFT model applied in this condition is complex for this reason. Except this reason, a graph method with estimating survival functions in Kaplan-Meier estimator would not derive the assumptions of proportional hazard when the Cox's model contains several time dependent covari-

ates. Hence, we tend to test this situation using the coding **cox.zph** in R and find a fit model of Cox's regression.

We can choose the best model for Cox's PH model with AIC criterion first of all. Since we are interested in impact of D-penicillamine and bilirubin to lifetime of patients, the two covariates must be contained in the last fitted model. Then we can use the code **stepAIC** in R for each first power covariate, and hence, the fit model is table I in Appendix II. Form the table I the test statistics of albumin and sex are 10.9861 and 5.8929 with p-values of 0.0009 and 0.0152, respectively, and the global test statistic is $\mathcal{X}^2 = 26.2377$ with a p-value of 0.0158. That means there exists some effects of nonlinearity for the above model and the global might not be suitable for the assumptions of Cox's PH model. So, we consider the martingale residual used for discovering the correct functional form for a predictor with covariates albumin and sex.

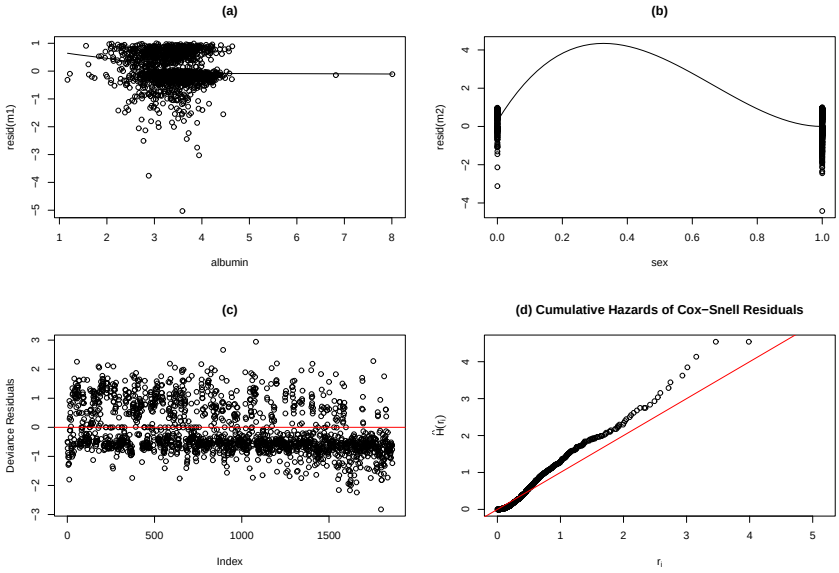

Figure 2: Diagnosis Tools in survival analysis on various residuals.

The covariate albumin might have the relationship of second power or log transformation with the Cox's PH model for this data and the covariate sex would not exist a linear effect obviously where the plots are printing in (a) and (b) of figure 2 and the Wald of the test of the hypothesis of no sex effect ($H_0 : \beta_f = 0$) is 10.9861 with a p-value of 0.0009. The results from the stratified model in this case are quite close to those obtained in the unstratified model. So we would like to stratify on the covariate sex and make the second power transformation to albumin. Then, the last fit model is

$$\hat{h}(t|\mathbf{Z}(t)) = h_0(t) \exp\{\mathbf{Z}(t)'\hat{\beta}\} \tag{3}$$

where

$$\begin{aligned}\mathbf{Z}(t)'\hat{\beta} &= -0.663Z_1 + 0.0413(bilirubin) + 0.0226(age) + 0.2540Z_2 + 0.5037Z_3 \\ &\quad +0.0018(SGOT) + 0.2799Z_4 + 0.0758(pt) - 0.0962(\log(platelets)) \\ &\quad -0.9755(albumin) + 0.0919(albumin)^2 - 0.0712(\log(alkaline)) \\ &\quad 0.1165Z_5\end{aligned}$$

and Z_1 is 1 if the patients take D-penicillamine and 0 if the patients take placebo; Z_2 is 1 if edema present without diuretics or edema resolved by diuretics and 0 if the other situation; Z_3 is 1 if edema despite diuretic therapy and 0 if the other condition; Z_4 is 1 if the presence of hepatomegaly for the individuals and 0 otherwise; Z_5 is 1 if the presence of spiders and 0 otherwise.

There is the deviance residual, which is a normalized transform of martingale residual, identifying poorly predicted individuals plotted in (c) of figure 2. The deviance residuals are all near to zero and there dose not exist outlier or any pattern in this plot, such as a sector and so on. Furthermore, all of points are uniformly scattered between -3 and 3. Therefore, the assumptions of Cox's PH are fitted in with model (3), that is, the fit model is suitable for Cox's PH model. For the other way, we are also have to plot the cumulative hazards function of Cox-Snell residuals provided a way of checking goodness of fit in (d) of figure 2. If the Cox's model fits the data, the plot should follow the 45^o line. Although the points are mostly up the red line (45^o line), we would still say that it follows the straight line with the slope of 1. So the plot suggests that this model dose not fit too badly. That means the Cox model fits the PBC data. Finally, we can use the code **cox.zph** in R to get a chi-square test statistic that provides a useful value to test if the Cox's model is fit for the PBC data. So the test statistic is equal to $\mathcal{X}^2 = 20.5961$ with a p-value of $0.0813 > 0.05$. That is, we do not have enough evidence to say that the assumptions of proportional hazards is not conformed with PBC data. And the detailed output form R programming is listed in table IV of Appendix II.

No matter what methods for testing various residuals, all of them show that the Cox's model is fit for PBC data and the best model is model (3) obtained by AIC criterion. Then, we can investigate the impact of D-penicillamine and bilirubin to lifetime of patients and obtain the information which covariate has significantly affect the patients' survival time by this way. Moreover, we will discuss several circumstances what we can know from the model (3).

4. Conclusion

By the AIC criterion, we have the fit Cox's PH model is in equation (3) based on the stratification of female. Then the below table illustrates convariates of consideration about the model (3).

Covariate	Est. Value	exp(Est. Value)	Std. Error	z	p
Z_1(Drug)	-0.0663	0.936	0.0792	-0.836	4.0e-01
bilirubin	0.0413	1.042	0.0075	5.490	4.0e-08
age in years	0.0266	1.027	0.0042	6.361	2.0e-10
Z_2(edema=0.5)	0.2540	1.289	0.0990	2.566	1.0e-02
Z_3(edema=1)	0.5034	1.655	0.1361	3.701	2.2e-04
SGOT	0.0018	1.002	0.0004	4.470	7.8e-06
Z_4(hepatomegaly)	0.2799	1.323	0.0932	3.004	2.7e-03
pt	0.0758	1.079	0.0258	2.938	3.3e-03
log(platelet)	-0.0962	0.908	0.0394	-2.440	1.5e-02
albumin	-0.9755	0.377	0.4603	-2.119	3.4e-02
$albumin^2$	0.0919	1.096	0.0670	1.313	1.9e-01
log(alkaline)	-0.0712	0.931	0.0374	-1.902	5.7e-02
Z_5(spider)	0.1165	1.124	0.0877	1.328	1.8e-01

In the above table, the p-values of covariates drug, $albumin^2$ and spider are 0.4, 0.19 and 0.18 and all of them are less than 0.05. That means we do not reject the test $H_0 : \beta_i = 0,\ i = 1, 11, 13$ at significant level 0.05. Thus, the three covariates do not exist significant effect obviously for the patients of survival time. However, those covariates added into Cox's PH model would make the value of log-likelihood corresponding to a fitted model decreasing clearly in the AIC argument and we are interested in the impact of D-penicillamine (drug= 1) to lifetime of patients with PBC data. Therefore we would still consider their effect for the patients' lifetime in model (3).

Next, the relative risk for Cox's PH model is considered and its defines as

$$\frac{\hat{h}(t|Z_i = 1)}{\hat{h}(t|Z_i = 0)} = \frac{h_0(t) \exp\left\{\hat{\beta}_i\right\}}{h_0(t)} = \exp\left\{\hat{\beta}_i\right\}.$$

So the relative risk for the patients using D-penicillamine as compared to using placebo is $\exp\{-0.0663\} = 0.936$ and its 95% confidence interval is $(0.8013, 1.0930)$, that is, the patients using D-penicillamine would have 0.936 times death risk as using placebo and we have 95% confidence level to say that patients given a placebo are between 0.9 and 1.2 times more likely to death than patients given D-penicillamine. Although the effect of drug is not significantly, patients given D-penicillamine would have longer lifetime than patients given placebo. Similarly, if bilirubin is equal to $x + 1$ and x then its relative risk is equal to $\exp\{0.0413\} = 1.042$ which means that if the bilirubin increases one milliliter then the individual's death rate would increase 1.042 times. And the normal bilirubin rage is between 0.2 and 1.2 milliliter in the medical science. Consequently, the bilirubin is larger than 20 milliliter then the individual would fall sick of serious illness, such as Liver Cirrhosis (L.C. for short). Thus, if the patient has high value of bilirubin then whether given D-

penicillamine or not would have a small inference even if non-impact to lifetime of patients.

On the other index, the relative risk for $Z_2 = 1$ and $Z_3 = 0$ as compared to $Z_2 = 0$ and $Z_3 = 0$ is $\exp\{0.2540\} = 1.289$ and its corresponding 95% confidence interval is (1.062, 1.565). Also, the relative risk for $Z_2 = 0$ and $Z_3 = 1$ as compared to $Z_2 = 0$ and $Z_3 = 0$ is $\exp\{0.5037\} = 1.655$ with 95% confidence interval (1.267, 2.161). In other words, the edema presents without diuretics or edema resolved by diuretics for patients and edema despite diuretic therapy have between 1.062 and 1.565 times and between 1.267 and 2.161 times more likely to death than that patients have no edema and no diuretic therapy for edema, respectively. So, the presence of edema for patients would have shorter lifetime with respect to no edema occurrence.

By the same argument, the presence of hepatomegaly for patients would have between 1.102 and 1.588 times as compared to none of hepatomegaly; the presence of spiders for patients has between 0.946 and 1.334 times as compared to none of spiders; and albumin increased one gm leads to death risk reduces between 0.153 and 0.929 times than before increasing, etc...Also, platelets are usual between 150 and 450 thousands per UL and patients fell L.C. would decrease the amount of platelets in clinical study. Furthermore, if the index of albumin is less than 2.5 gm/dl then patients would fall sick L.C. and have the presence of hepatomegaly and ascites at the same time. Finally, alkaline means alkaline phosphatase and its normal index is between 43 and 122 iu/L and the illness, L.C., might make this index descended. From the ahead statement, the estimators of covariates estimated with Cox's PH model are the same as the true situations. Thus, three indexes would increase the lifetime of patients when they are increasing. Also, the covariates Z_1 (drug), log(platelet), albumin and alkaline would make death rate decreasing, i.e., these covariates would cause longer survival time. And the rest would reduce patients' lifetime. The connected outputs are in table III of Appendix II.

5. Discussion

We could know that the covariates Z_1, log(platelets), log(alkalne), Z_4 and Z_5 contained in our last model are not sufficiently effects for the lifetime of patients and disease's stage is sufficiently with the criterion AIC in Zhu M. and Fan G. and Fan G. (2006). However, there has some different form our results, that is, the covariate, disease's stage, was not contained in our last model and we intended to say that log(platelets), log(alkalne) and Z_4 have sufficiently influences in table III of Appendix II. Although there are more covariates choosed with Cox's PH model in this project than in Zhu M. and Fan G. (2006), the important covariates except the disease's stage which influence the lifetime of patients are still contained in our last model. This shows we might omit some imformations of the PBC data not to consider and we could adjust by transforming covaraites.

Furthermore, we do not apply the AFT model to this data concluded time dependent covariates. Here has a similar method, step stress accelerated degradation tests, which is useful to analysis time dependent covariates in reliability. Second, there are many missing observations in this PBC data and we remove the whole rows includes missing data. But we could consider an argumented inverse probability weighted (AIPW) estimator of Wang and Chen (2001) that avoids its stronger censorship condition through inverse probabilty weighting. Upon inserting the missingness probabilities or consistent estimates thereof, this estimator can be calculated using an EM procendure. Thus, we could consider the step stress and Cox's model estimated parameters with using AIPW estimator to apply a data containing imcomplete data (or missing data) and more than one time dependent covariate. By those parametric and semi-parametric ways, we could find which model is the best model to fit a data.

References

Liao, C.M. and Tseng, S.T. (2006). Optimal Design for Step-Stress Accelerated Degradation Tests. IEEE TRANSACTIONS ON RELIABILITY, VOL. 55, 59-66.

Elizabeth J. Malloy, Donna Spiegelman , Ellen A. Eisen (2009). Comparing measures of model selection for penalized splines in Cox models. Computational Statistics and Data Analysis 53, 2605-2616.

John D. and Ross L. (2002) The Statistical Analysis of Failure Time Data , Wiley.

Terry M. and Patricia M. (2000) Modeling Survival Data: Extending the Cox Model, Springer.

Zhu M. and Fan G. (2006) Variable Selection in Cox's Proportional Hazards Model Using a Parallel Genetic Algorithm.

Appendix I : Data Introduction

Sequential PBC

The variables in the data set are :

case number;

number of days between resistration and the earlier death, transplantation, or study analysis time;

status: 0=alive, 1=transplanted, 2=death;

drug: 1=D-penicillamine, 0=placebo;

age in days, at restration;

sex: 0=male, 1=female;

day: number of days between enrollment and this visit date; remaining values on the line of data refer to this visit;

presence of ascites: 0=no, 1=yes;

presence of hepatomegaly: 0=no, 1=yes;

presence of spiders: 0=no, 1=yes;

presence of edema: 0=no edema and no diuretic therapy for edema, 0.5=edema present without diuretics, or edema resolved by diuretics; 1=edema despite diuretic therapy;

serum bilirubin in mg/dl;

albumin in gm/dl;

alkaline phosphatase in U/liter;

SGOT in U/ml;

platelets per cubic ml/1000;

prothrombin time in seconds; and

histologic stage of disease.

Appendix II : Tables and Figures

Table I: model selection with AIC criterion

Covariate	ρ	chisq	p
Z_1(drug=1)	-0.0133	0.1194	0.7297
bilirubin	0.0450	1.1559	0.2823
albumin	-0.1264	10.9861	0.0009
age	0.0099	0.0697	0.7918
sex(female)	-0.0897	5.8929	0.0152
edema(0.5)	0.0116	0.0930	0.7604
edema(1)	0.0207	0.2928	0.5884
SGOT	-0.0234	0.2066	0.6494
hepatomegaly(1)	-0.0136	0.1251	0.7235
pt	-0.0117	0.0618	0.8037
log(platelets)	0.0385	1.0477	0.3060
log(alkaline)	0.0418	1.1281	0.2882
spiders(1)	-0.0279	0.5015	0.4788
GLOBAL	NA	26.2377	0.0158

Table II: AIC values of transformating albumin and startifying sex

Transformation	df	AIC
albumin only	12	6143.961
albumin and albumin^2	13	6144.62
albumin^{-1} only	12	6145.199
log(albumin) only	12	6143.38
sqrt(albumin) only	12	6143.434

Table II shows that the covariates entered into the Cox's model $h(t|\mathbf{Z}(t)) = h_0(t)\exp\{\mathbf{Z}'(t)\beta\}$ *where* $\mathbf{Z}(t)$ *means the covariates in table I. With AIC criteria, we should choose the condition with minimizes AIC value, i.e., add the covariate log(albumin) into the above Cox's model. But the chi-square test statistic is 8.2594 with a p-value of 0.0041<0.05, this shows that the covariate log(albumin) does not conform to the assumptions of Cox's PH model. However, the likelihood ratio test statistic between the two models, addding albumin, albumin*^{2} *and log(albumin) only, is 2 and its p-value is 0.1573>0.05. This means that there dose not significant difference between these models. Thus, we would attend to choose the situation that a model contains the covariates albumin and albumin*^{2}*.*

Table III: The last fit model with Cox's PH model

Covaraite	Est. Value	Std. Error	z	p	lower .95	upper .95
Z_1(drug=1)	-0.0663	0.0792	-0.836	4.0e-01	0.801	1.093
bilirubin	0.0413	0.0075	5.490	4.0e-08	1.027	1.058
age in years	0.0266	0.0042	6.361	2.0e-10	1.019	1.035
Z_2(edema=0.5)	0.2540	0.0990	2.566	1.0e-02	1.062	1.565
Z_3(edema=1)	0.5037	0.1361	3.701	2.2e-04	1.267	2.161
SGOT	0.0018	0.0004	4.470	7.8e-06	1.001	1.003
Z_4(hepatomegaly)	0.2799	0.0932	3.004	2.7e-03	1.102	1.588
pt	0.0758	0.0258	2.938	3.3e-03	1.026	1.135
log(platelets)	-0.0962	0.0394	-2.440	1.5e-02	0.841	0.981
albumin	-0.9755	0.4603	-2.119	3.4e-02	0.153	0.929
albumin^2	0.0919	0.0670	1.313	1.9e-01	0.956	1.257
log(alkaline)	-0.0712	0.0374	-1.902	5.7e-02	0.865	1.002
Z_5(spider=1)	0.1165	0.0877	1.328	1.8e-01	0.946	1.334

The lower .95 and upper .95 are the lower bound and upper bound of 95% *confidence interval for relative risk, respectively. The z values in table III obtain from the Wald's test and p's are the p-values with respect to z that are using to test the null hypotheses $H_0 : \beta_i = 0,\ \forall i$. Also, the red words show that β_i's are not equal to 0 obviously at significant level 0.05.*

Table IV: The last fit model for checking the PH assumptions

Covariate	ρ	chisq	p
Z_1(drug=1)	-0.0090	0.0561	0.8128
bilirubin	0.0499	1.4163	0.2340
age in years	0.0237	0.3985	0.5278
Z_2(edema=0.5)	0.0254	0.4500	0.5024
Z_3(edema=1)	0.0305	0.6223	0.4302
SGOT	-0.0333	0.4189	0.5175
Z_4(hepatomegaly)	-0.0093	0.0598	0.8069
pt	-0.0153	0.1045	0.7465
log(platelets)	0.0496	1.7702	0.1834
albumin	-0.0846	1.9694	0.1605
albumin^2	0.0537	0.7360	0.3910
log(alkaline)	0.0431	1.2079	0.2718
Z_5(spider=1)	-0.0164	0.1777	0.6733
GOLBAL	NA	20.5961	0.0813

The p-value of GOLBAL is 0.0813>0.05, that is, we do not reject the H_0: the PH assumptions hold. So we could believe that this model conform to the PH assumptions and using Cox's PH model to fit the PBC data is right.

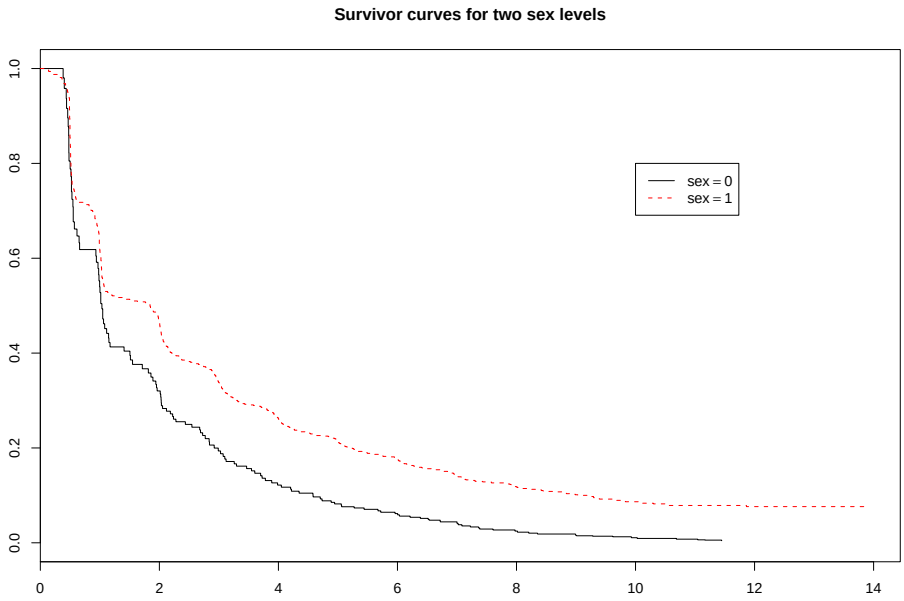

Figure 3: The survival curvs fitted with stratification sex.

This figure is the survivavor curvs of each stratification of sex fitted with KM estimator of the last fit Cox's model where the red line means the survival probabilty of female for the patients and the black line means survival probability of male for the individuals. Form this figure, we could know that the lifetime of female for the patients is longer than males. On the other hand, there has larger death risk for males then females.

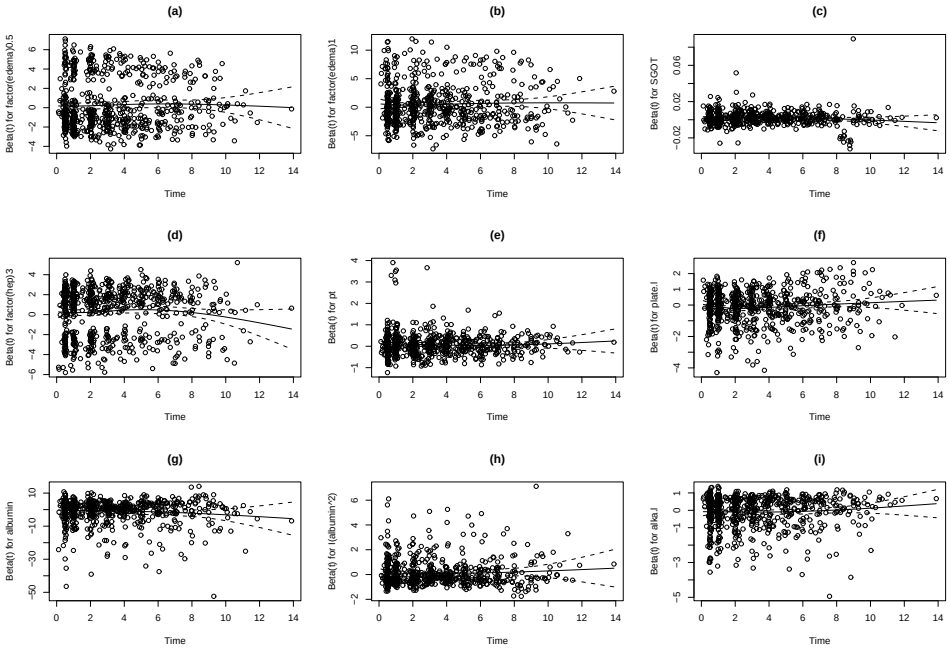

Figure 4: Testing the time dependent for all discrete covariates.

We could know whether the covariates were time dependent or not from these figures. In plot (d), the line dose not seem to be a horizontal line, but this condition is not very clearly and there has a few onservations at the back time. So, we also could acept it is a horizontal line. And the rest plots were arising horizontal lines for every different covariate. Although we make sure that several covariates in the last fit Cox's model are time dependent covariates, we could still see that all covariates contained in the last fit Cox's model do not exist time dependent obviously. However, it is not proper to test time dependent covariate by a graph, and hence, the above figures would be used as a consultation.

Appendix III : Codes for Programming R

```
library(MASS)
library(survival)
library(nlme)
status = ifelse(status == 2, 1, 0)
## Transformating the time in days to in years
age = age / 365.24
begin = begin / 365.24 ; endtime = endtime / 365.24
S = Surv(begin, endtime, status)
alka.l = log(alka)
plate.l = log(plate)

## Fitting the Kaplan-Meier estimator for drug and sex
fit.drug = survfit(S ~ factor(drug))
fit.sex = survfit(S ~ factor(sex))
##Plotting the KM survival curv
plot(fit.drug[1] $ time, fit.drug[1] $ surv, type = "s", xlab = "Time(year)",
     ylab = "S(t)", main = "Survival curves for the covariate drug")
##Plotting the cumulative hazards figures
##exchange fit.drug with fit.sex and then the figures about sex are ploted
plot(fit.drug[1] $ time, log(-log(fit.drug[1] $ surv)), type = "s", xlab =
"Time(year)", ylab = "log-log S(t)", main = "Proportional hazard testing for drug")

##Log-rank test
surv.time = num / 365.24
S = Surv(surv.time, fustat)
test.drug = survdiff(S ~ factor(drug))
test.sex = survdiff(S ~ factor(sex))

##Model selection with AIC criterion
Scope = list(upper = ~ (factor(drug) + bilirubin + age + factor(sex) + factor(asci)
             + factor(hep) + factor(spider) + factor(edema) + albumin + alka.l +
             SGOT + plate.l + pt + factor(stage)), lower = ~(factor(drug) +
             bilirubin))
pm_0 = coxph(S ~ factor(drug) + bilirubin)
pm_f = stepAIC(pm_0, Scope, trace = F)
## pm_f is the fit model for this data

m1 = coxph(S ~ factor(drug) + bilirubin + age + factor(sex) + factor(edema) +
           SGOT + factor(hep) + pt + plate.l + alka.l + factor(spider))
m2 = coxph(S ~ factor(drug) + bilirubin + age + factor(edema) +
```

```
           SGOT + factor(hep) + pt + plate.l + albumin + alka.l + factor(spider))
##Computing each model's value of log likelihood
extractAIC(m1)[2]
extractAIC(m2)[2]
##Plotting the survival curv estimated by KM estimator
##for the stratification sex.
plot(survfit(mi), lty = 1 : 2, col = 1 : 2, main =
     "Survivor curves for two sex levels")

##Plotting the figure to show if the covariates are time dependent.
fit.cox = cox.zph(mi, transform = 'identity')
plot(fit.cox[4], main = "(a)") ; plot(fit.cox[5], main = "(b)")
plot(fit.cox[6], main = "(c)") ; plot(fit.cox[7], main = "(d)")
plot(fit.cox[8], main = "(e)") ; plot(fit.cox[9], main = "(f)")
plot(fit.cox[10], main = "(g)") ; plot(fit.cox[11], main = "(h)")
plot(fit.cox[12], main = "(i)")
```