Survival Analysis for Primary Biliary Cirrhosis Data by using SAS

Preprint · February 2018

DOI: 10.13140/RG.2.2.23354.90569

1 author:

Mayooran Thevaraja
University of Jaffna
16 PUBLICATIONS **34** CITATIONS

SEE PROFILE

Survival Analysis for Primary Biliary Cirrhosis Data by using SAS

A Research Project Report Submitted By

Mayooran, Thevaraja

this project report is submitted in partial fulfilment of the requirement for the STAT654 (Statistical Methods for Biostat/Survival Analysis) Course

Department of Mathematics and Statistics,
Minnesota State University, Mankato, USA,
February 2018

Survival Analysis for Primary Biliary Cirrhosis Data by using SAS

Mayooran, Thevaraja
Department of Mathematics and Statistics,
Minnesota State University, Mankato, USA
Email: thevaraja.mayooran@mnsu.edu

Abstract

This research works discussed of the Kaplan-Meier estimator and a real data, the sequential primary biliary cirrhosis collected in Mayo clinic, which holds several time dependent covariates and the observations is measured repeatedly fits for the Cox proportional hazard models because the Cox proportional hazard model is the most popular method for survival analysis. Also, by employing the model fit criterion to figure out a suitable model for the real data and the criterion considered were the Akaike information criterion(AIC) for model validation. We find some factors, which have time to event, are agreed the assumptions of Cox proportional hazard model. So, the Cox PH model is proper to the data and the lifetime of patients would be related strongly to the covariates bilirubin, age, presence of edema, SGOT, presence of hepatomegaly, pro-thrombin time, platelets, albumin and alkaline.

Keywords: *Cox proportional hazard model; Kaplan-Meier estimate; AIC;*

Contents

1. Introduction.

Bile is a fluid formed in human liver that works in the digestive process and will help human remove human blood from red blood, cholesterol and toxins. Biliary cirrhosis is the most commonly the autoimmune disease that is in its cells, in the case of bile ducts. As the water cycle increases, harmful substances can accumulate. This can lead to the liver (this is cirrhosis). Among other things, the sufferer may experience abdominal pain, diarrhea and, finally, the liver. The key to biliary cirrhosis is a major risk for cancer. This illustration uses data from a randomized controlled trial of D-penicillamine (DPCA) for the treatment of primary biliary cirrhosis. A total of n=312 consenting subjects were enrolled and randomized to either active treatment or placebo-control (this group received standard care). Time zero is date of diagnosis and initiation of treatment. Study participants were followed to event of end-stage liver disease or censoring. Thus, these are an example of "right" censored data. Over the approximate 10 years of follow-up, 125 events of death (40%) were seen.

2. Literature Review

Already the several scholars in various locations all over the world have done this statistical analysis related medical data survival analysis. The general idea of survival analysis is to create what are known as survival models or survival curves and fit the cox regression model. A survival model or curve is a way of representing how long a patient with a given disease or condition will survive (Balakrishnan etl 2004). While there are quite a few different methods of doing this, the overall idea is to create a model showing the probability of surviving a certain length of time for each type of treatment. These models are based off of empirical data collected by medical researchers of people who are actually being treated using either the new or the old method (Guo 1991). The most simple versions work based solely on how long each individual survived, but more complicated versions also exist, which can take into account a number of other variables (medical history, pre-existing diseases, lifestyle, ethnicity, age, gender, etc.), which are considered to be important in the medical world (Selvin 2008). In Minnesota, C.W. Woodall (2005) and et al studied survival analysis to a large-scale forest inventory for assessment of tree mortality in Minnesota. David S. Knopman, MD, Department of Neurology, Mayo Clinic, Rochester, studied about investigate the relationship between features and definitions of vascular dementia and survival.

3. Data preparation and exploration

3.1 Structure of the data

The data is from the Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver conducted between 1974 and 1984. A total of 424 PBC patients, referred to Mayo Clinic during that ten-year interval, met eligibility criteria for the randomized placebo-controlled trial of the drug D-

penicillamine. The first 312 cases on the data set aside and determine randomized and hold full data. This data has many laboratory studies, but only 312 patients, one of them left 140 and the rest left, at least 9: 1 (female and male) as databases. The other 112 cases have not been involved in the trial, but they have agreed to have the same degree of responsibility as to the extent to which they will apply for survival. Sixty of those things depend on shortly after discovery, so this data is in 106 cases and 312 people. The purpose of this study is to determine the effects of D-penicillamine and bilirubin for patients living in Primary Biliary Cirrhosis (PBC). The data record gives to drug addicts, diabetes and enrollment, sex, presence of ascites, fluids, spiders and edema, bilirubin blood, albumin, alkaline phosphatase, serum glutamic-oxaloacetic transaminase (SGOT), cubic platelets, prothrombin periods and the history of the disease, etc.

3.2 Data exploration

Any serious endeavor into data analysis should begin with data exploration, in which the researcher becomes familiar with the distributions and typical values of each variable individually, as well as relationships between pairs or sets of variables. This database is the entry and the first PBC data entry, as well as the text for the search for each person's study. Databases based on the data found in Murtaugh PA are found. Dickson ER. Van Dam GM. Malinchoc Mr. Grambsch PM. Langworthy AL. Gips CH. "The first biliary cirrhosis: The survival policy depends on the visit repeatedly." Hepatology. 20 (1.1): 126-34, 1994. The basic PBC databases have only the basic configuration process of the search engineer. This data base contains several laboratory detections, but it is only the first 312. Some databases on this file are different from the original PBC files, for example, data errors and prothrombin times and years that are detected during analytics, at work time surveys and datasets. (These two points are discussed in Fleming and Harrington, 4.6.7). Another important factor is that there are many other factors that affect patients when the data collection is collected. One "character" of the data should speak special. The last vision before loss or liver has much to cost covariates than other data lines. The first treatment of these patients is six months, 1 year, and every year later. In these procedures are the visits to the laboratory for a large battery of test. The "Extra" Visitor, which is usually due to the health status, is likely to have been a major part of the homework. Such unexpected rules can get information, and damage the lost costs "(MCAR or MACs) that are considered in the research. From the first published results on the mark Mayo PBC, however, usually contains five percent of the compounds, namely, age, bilirubin, albumin, prothrombin time, and diarrhea.

4. Methodology

4.1.1 The Kaplan-Meier estimator of the survival function

The Kaplan Meier survival function estimator is calculated as:

$$\hat{S}_t = \prod_{t_i \leq 1}^{n} \frac{n_i - d_i}{n_i}$$

where n_i is the number of subjects at risk and d_i is the number of subjects who fail, both at time t_i. Thus, each term in the product is the conditional probability of survival beyond time t_i, meaning the probability of surviving beyond time t_i, given the subject has survived up to time t_i. The survival function estimate of the the unconditional probability of survival beyond time t (the probability of survival beyond time t from the onset of risk) is then obtained by multiplying together these conditional probabilities up to time t together.

4.1.2 Nelson-Aalen estimator of the cumulative hazard function

The simple relationship with the survival function and hazard function,\ S(t) = e^{$-H(t)$}, the cumulative hazard function can be used to estimate the survival function. The Nelson-Aalen estimator is a non-parametric estimator of the cumulative hazard function and is given by:

$$\widehat{H}_t = \sum_{t_i} \frac{d_i}{n_i}$$

where d_i is the number who failed out of n_i at risk in interval t_i. The estimator is calculated, then, by summing the proportion of those at risk who failed in each interval up to time t.

4.2. Comparing survival functions using nonparametric tests

4.2.1 Tests of equality of the survival function

In the output we find three Chi-square based tests of the equality of the survival function over strata, which support our suspicion that survival differs between genders. The calculation of the statistic for the nonparametric "Log-Rank" and "Wilcoxon" tests is given by:

$$Q = \frac{\left[\sum_{i=1}^{m} w_j (d_{ij} - \hat{e}_{ij})\right]^2}{\sum_{i=1}^{m} w_j^2 \hat{v}_{ij}}$$

where d_{ij} is the observed number of failures in stratum i at time t_j, $\hat{e}_{ij}$ is the expected number of failures in stratum i at time t_j $\hat{v}_{ij}$ is the estimator of the variance of d_{ij}, and w_i is the weight of the difference at time t_j (see Hosmer and Lemeshow(2008) for formulas for $\hat{e}_{ij}$and $\hat{v}_{ij}$). In a nutshell, these statistics sum the weighted differences between the observed number of failures and the expected number of failures for each stratum at each timepoint, assuming the same survival function of each stratum. In other words, if all strata have the same survival function, then we expect the same proportion to die in each interval. If these proportions systematically differ among strata across time, then the Q statistic will be large and the null hypothesis of no difference among strata is more likely to be rejected. The log-rank and Wilcoxon tests in the output table differ in the weights w_j used. The log-rank or Mantel-Haenzel test uses w_j=1, so differences at all time intervals are weighted equally. The Wilcoxon test uses w_j=n_j, so that differences are weighted by the number at risk at time t_j, thus giving more weight to differences that occur earlier in follow-up time. The "-2Log(LR)" likelihood ratio test is a parametric test assuming exponentially distributed survival times and will not be further discussed in this nonparametric section.

4.2.2 Nonparametric estimation of the hazard function

Standard nonparametric techniques do not typically estimate the hazard function directly. However, we can still get an idea of the hazard rate using a graph of the kernel-smoothed estimate. As the hazard function $h(t)$ is the derivative of the cumulative hazard function $H(t)$, we can roughly estimate the rate of change in $H(t)$ by taking successive differences in $\widehat{H}(t)$ between adjacent time points, $\Delta\widehat{H}(t) = \widehat{H}(t_j) - \widehat{H}(t_{j-1})$. SAS computes differences in the Nelson-Aalen estimate of $H(t)$. We generally expect the hazard rate to change smoothly (if it changes) over time, rather than jump around haphazardly. To accomplish this smoothing, the hazard function estimate at any time interval is a weighted average of differences within a window of time that includes many differences, known as the bandwidth. Widening the bandwidth smooths the function by averaging more differences together. However, widening will also mask changes in the hazard function as local changes in the hazard function are drowned out by the larger number of values that are being averaged together.

4.3. The Cox proportional hazards regression mode

4.3.1 Estimating the hazard function

Whereas with non-parametric methods we are typically studying the survival function, with regression methods we examine the hazard function, $h(t)$. The hazard function for a particular time interval gives the probability that the subject will fail in that interval, given that the subject has not failed up to that point in time. The hazard rate can also be interpreted as the rate at which

failures occur at that point in time, or the rate at which risk is accumulated, an interpretation that coincides with the fact that the hazard rate is the derivative of the cumulative hazard function, $H(t)$. In regression models for survival analysis, we attempt to estimate parameters which describe the relationship between our predictors and the hazard rate. We would like to allow parameters, the β s, to take on any value, while still preserving the non-negative nature of the hazard rate. A common way to address both issues is to parameterize the hazard function as:

$$h(t|x) = exp(\beta_0 + \beta_1 x)$$

In this parameterization, $h(t|x)$ is constrained to be strictly positive, as the exponential function always evaluates to positive, while β_0 and β_1 are allowed to take on any value. Notice, however, that t does not appear in the formula for the hazard function, thus implying that in this parameterization, we do not model the hazard rate's dependence on time. A complete description of the hazard rate's relationship with time would require that the functional form of this relationship be parameterized somehow (for example, one could assume that the hazard rate has an exponential relationship with time). However, in many settings, we are much less interested in modeling the hazard rate's relationship with time and are more interested in its dependence on other variables, such as experimental treatment or age. For such studies, a semi-parametric model, in which we estimate regression parameters as covariate effects but ignore (leave unspecified) the dependence on time, is appropriate.

4.3.2 The Cox proportional hazards model

We can remove the dependence of the hazard rate on time by expressing the hazard rate as a product of $h_0(t)$, a baseline hazard rate which describes the hazard rates dependence on time alone, and $r(x, \beta_x)$, which describes the hazard rates dependence on the other x

Covariates: $h(t) = h_0(t)r(x, \beta_x)$

In this parameterization, $h(t)$ will equal $h_0(t)$ when $r(x, \beta_x)$=1. It is intuitively appealing to let $r(x, \beta_x)$=1 when all $x = 0$, thus making the baseline hazard rate, $h_0(t)$ equivalent to a regression intercept. Above, we discussed that expressing the hazard rate's dependence on its covariates as an exponential function conveniently allows the regression coefficients to take on any value while still constraining the hazard rate to be positive. The exponential function is also equal to 1 when its argument is equal to 0. We will thus let $r(x, \beta_x) = exp(x\beta_x)$, and the hazard function will be given by: $h(t) = h_0(t)exp(x\beta_x)$

This parameterization forms the *Cox proportional hazards model*. It is called the proportional hazards model because the ratio of hazard rates between two groups with fixed covariates will stay constant over time in this model. For example, the hazard rate when time *t*
when $x = x_1$ would then be $h(t|x_1) = h_0(t)exp(x_1\beta_x)$and at time *t* when $x = x_2$ would be $h(t|x_2) = h_0(t)exp(x_2\beta_x)$ The covariate effect of *x*, then is the ratio between these two hazard rates, or a hazard ratio(HR):

$$HR = \frac{h(t|x_1)}{h(t|x_2)}) = \frac{h_0(t)exp(x_1\beta_x)}{h_0(t)exp(x_2\beta_x)}$$

Notice that the baseline hazard rate, $h_0(t)$ is cancelled out, and that the hazard rate does not depend on time *t*:

$$HR = exp(\beta_x(x_2 - x_1))$$

The hazard rate *HR* will thus stay constant over time with fixed covariates. Because of this parameterization, covariate effects are multiplicative rather than additive and are expressed as hazard ratios, rather than hazard differences. As we see above, one of the great advantages of the Cox model is that estimating predictor effects does not depend on making assumptions about the form of the baseline hazard function, $h_0(t)$, which can be left unspecified. Instead, we need only assume that whatever the baseline hazard function is, covariate effects multiplicatively shift the hazard function and these multiplicative shifts are constant over time. Cox models are typically fitted by maximum likelihood methods, which estimate the regression parameters that maximize the probability of observing the given set of survival times. We can similarly calculate the joint probability of observing each of the *n* subject's failure times, or the likelihood of the failure times, as a function of the regression parameters, β, given the subject's covariates values x_j:

$$\mathrm{L}(\beta) = \prod_{i=1}^{n} \frac{exp(x_j\beta_x)}{\sum exp(x_i\beta_x)}$$

where R_j is the set of subjects still at risk at time t_j. Maximum likelihood methods attempt to find the β values that maximize this likelihood, that is, the regression parameters that yield the maximum joint probability of observing the set of failure times with the associated set of covariate values. Because this likelihood ignores any assumptions made about the baseline hazard function, it is actually a partial likelihood, not a full likelihood, but the resulting β have the same distributional properties as those derived from the full likelihood. Finally we will Fit a simple Cox regression model for this data and discuss some results.

5. Result and Discussions

Survival analysis often begins with examination of the overall survival experience through non-parametric methods, such as Kaplan-Meier (product-limit) and life-table estimators of the survival function. Non-parametric methods are appealing because no assumption of the shape of the

survivor function nor of the hazard function need be made. However, nonparametric methods do not model the hazard rate directly nor do they estimate the magnitude of the effects of covariates. So, in this research we are going to use `proc lifetest` function in SAS, at a minimum `proc lifetest` requires specification of a failure time variable, here `futime`, on the time statement.

Product-Limit Survival Estimates

futime	Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.00	312	0	1.0000	0	0	0	312
41.00	312	1	0.9968	0.00321	0.00320	1	311
51.00	311	1	0.9936	0.00641	0.00452	2	310
71.00	310	1	0.9904	0.00962	0.00552	3	309
77.00	309	1	0.9872	0.0128	0.00637	4	308
110.00	308	1	0.9840	0.0160	0.00711	5	307
130.00	307	1	0.9808	0.0192	0.00778	6	306
131.00	306	1	0.9776	0.0224	0.00838	7	305
140.00	305	1	0.9744	0.0256	0.00895	8	304
179.00	304	1	0.9712	0.0288	0.00948	9	303
186.00	303	1	0.9679	0.0321	0.00997	10	302
191.00	302	1	0.9647	0.0353	0.0104	11	301
198.00	301	1	0.9615	0.0385	0.0109	12	300
207.00	300	1	0.9583	0.0417	0.0113	13	299
216.00	299	1	0.9551	0.0449	0.0117	14	298
223.00	298	1	0.9519	0.0481	0.0121	15	297
264.00	.	.	.	.	.	16	296
264.00	297	2	0.9455	0.0545	0.0129	17	295
304.00	295	1	0.9423	0.0577	0.0132	18	294
321.00	294	1	0.9391	0.0609	0.0135	19	293
326.00	293	1	0.9359	0.0641	0.0139	20	292
334.00	292	1	0.9327	0.0673	0.0142	21	291
348.00	291	1	0.9295	0.0705	0.0145	22	290
388.00	290	1	0.9263	0.0737	0.0148	23	289

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
400.00		289	1	0.9231	0.0769	0.0151	24	288
460.00		288	1	0.9199	0.0801	0.0154	25	287
515.00		287	1	0.9167	0.0833	0.0156	26	286
533.00	*	286	0	.	.	.	26	285
549.00		285	1	0.9135	0.0865	0.0159	27	284
552.00		284	1	0.9102	0.0898	0.0162	28	283
597.00		283	1	0.9070	0.0930	0.0164	29	282
611.00		282	1	0.9038	0.0962	0.0167	30	281
673.00		281	1	0.9006	0.0994	0.0169	31	280
694.00		280	1	0.8974	0.1026	0.0172	32	279
708.00		279	1	0.8942	0.1058	0.0174	33	278
732.00	*	278	0	.	.	.	33	277
733.00		277	1	0.8909	0.1091	0.0177	34	276
737.00	*	276	0	.	.	.	34	275
750.00		275	1	0.8877	0.1123	0.0179	35	274
762.00		274	1	0.8844	0.1156	0.0181	36	273
769.00		273	1	0.8812	0.1188	0.0183	37	272
786.00		272	1	0.8780	0.1220	0.0186	38	271
788.00	*	271	0	.	.	.	38	270
790.00		270	1	0.8747	0.1253	0.0188	39	269
797.00		269	1	0.8715	0.1285	0.0190	40	268
799.00		268	1	0.8682	0.1318	0.0192	41	267
824.00		267	1	0.8650	0.1350	0.0194	42	266
837.00	*	266	0	.	.	.	42	265
839.00	*	265	0	.	.	.	42	264
850.00		264	1	0.8617	0.1383	0.0196	43	263
853.00		263	1	0.8584	0.1416	0.0198	44	262

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
859.00		262	1	0.8551	0.1449	0.0200	45	261
877.00	*	261	0	.	.	.	45	260
890.00		260	1	0.8518	0.1482	0.0202	46	259
901.00	*	259	0	.	.	.	46	258
904.00		258	1	0.8485	0.1515	0.0204	47	257
930.00		257	1	0.8452	0.1548	0.0205	48	256
939.00	*	256	0	.	.	.	48	255
943.00		255	1	0.8419	0.1581	0.0207	49	254
971.00		254	1	0.8386	0.1614	0.0209	50	253
974.00		253	1	0.8353	0.1647	0.0211	51	252
980.00		252	1	0.8320	0.1680	0.0213	52	251
994.00	*	251	0	.	.	.	52	250
999.00		250	1	0.8287	0.1713	0.0214	53	249
1000.00		249	1	0.8253	0.1747	0.0216	54	248
1012.00		248	1	0.8220	0.1780	0.0218	55	247
1030.00	*	247	0	.	.	.	55	246
1037.00		246	1	0.8187	0.1813	0.0219	56	245
1067.00	*	245	0	.	.	.	56	244
1077.00		244	1	0.8153	0.1847	0.0221	57	243
1080.00		243	1	0.8119	0.1881	0.0223	58	242
1083.00		242	1	0.8086	0.1914	0.0224	59	241
1084.00	*	241	0	.	.	.	59	240
1149.00	*	240	0	.	.	.	59	239
1152.00		239	1	0.8052	0.1948	0.0226	60	238
1153.00	*	238	0	.	.	.	60	237
1165.00		237	1	0.8018	0.1982	0.0227	61	236
1170.00		236	1	0.7984	0.2016	0.0229	62	235

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
1191.00		.	.	.	.	.	63	234
1191.00		235	2	0.7916	0.2084	0.0232	64	233
1212.00		233	1	0.7882	0.2118	0.0234	65	232
1216.00	*	.	0	.	.	.	65	231
1216.00	*	232	0	.	.	.	65	230
1217.00		230	1	0.7848	0.2152	0.0235	66	229
1230.00	*	229	0	.	.	.	66	228
1234.00	*	228	0	.	.	.	66	227
1235.00		227	1	0.7813	0.2187	0.0237	67	226
1250.00	*	226	0	.	.	.	67	225
1271.00	*	225	0	.	.	.	67	224
1293.00	*	224	0	.	.	.	67	223
1295.00	*	223	0	.	.	.	67	222
1297.00		222	1	0.7778	0.2222	0.0238	68	221
1300.00	*	221	0	.	.	.	68	220
1301.00	*	220	0	.	.	.	68	219
1302.00	*	219	0	.	.	.	68	218
1320.00	*	218	0	.	.	.	68	217
1321.00	*	217	0	.	.	.	68	216
1329.00	*	216	0	.	.	.	68	215
1349.00	*	215	0	.	.	.	68	214
1350.00		214	1	0.7742	0.2258	0.0240	69	213
1356.00		213	1	0.7705	0.2295	0.0241	70	212
1360.00		212	1	0.7669	0.2331	0.0243	71	211
1363.00	*	.	0	.	.	.	71	210
1363.00	*	211	0	.	.	.	71	209
1401.00	*	209	0	.	.	.	71	208

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
1408.00	*	208	0	.	.	.	71	207
1412.00	*	207	0	.	.	.	71	206
1413.00		206	1	0.7632	0.2368	0.0245	72	205
1418.00	*	205	0	.	.	.	72	204
1420.00	*	204	0	.	.	.	72	203
1427.00		203	1	0.7594	0.2406	0.0246	73	202
1433.00	*	202	0	.	.	.	73	201
1434.00		201	1	0.7556	0.2444	0.0248	74	200
1434.00	*	.	0	.	.	.	74	199
1435.00	*	199	0	.	.	.	74	198
1444.00		198	1	0.7518	0.2482	0.0250	75	197
1447.00	*	197	0	.	.	.	75	196
1455.00	*	196	0	.	.	.	75	195
1457.00	*	195	0	.	.	.	75	194
1481.00	*	194	0	.	.	.	75	193
1487.00		193	1	0.7479	0.2521	0.0251	76	192
1492.00		192	1	0.7440	0.2560	0.0253	77	191
1504.00	*	191	0	.	.	.	77	190
1525.00	*	190	0	.	.	.	77	189
1536.00		189	1	0.7401	0.2599	0.0255	78	188
1542.00	*	188	0	.	.	.	78	187
1558.00	*	187	0	.	.	.	78	186
1568.00	*	186	0	.	.	.	78	185
1569.00	*	185	0	.	.	.	78	184
1576.00		184	1	0.7361	0.2639	0.0256	79	183
1592.00	*	183	0	.	.	.	79	182
1614.00	*	.	0	.	.	.	79	181

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
1614.00	*	182	0	.	.	.	79	180
1615.00	*	180	0	.	.	.	79	179
1656.00	*	179	0	.	.	.	79	178
1657.00		178	1	0.7319	0.2681	0.0258	80	177
1666.00	*	177	0	.	.	.	80	176
1677.00	*	176	0	.	.	.	80	175
1682.00		175	1	0.7278	0.2722	0.0260	81	174
1690.00		.	.	.	.	.	82	173
1690.00		174	2	0.7194	0.2806	0.0264	83	172
1701.00	*	172	0	.	.	.	83	171
1702.00	*	171	0	.	.	.	83	170
1735.00	*	170	0	.	.	.	83	169
1741.00		169	1	0.7151	0.2849	0.0266	84	168
1765.00	*	168	0	.	.	.	84	167
1769.00	*	167	0	.	.	.	84	166
1770.00	*	166	0	.	.	.	84	165
1776.00	*	165	0	.	.	.	84	164
1783.00	*	164	0	.	.	.	84	163
1785.00	*	163	0	.	.	.	84	162
1786.00		162	1	0.7107	0.2893	0.0268	85	161
1790.00	*	161	0	.	.	.	85	160
1810.00	*	160	0	.	.	.	85	159
1827.00		159	1	0.7063	0.2937	0.0270	86	158
1831.00	*	158	0	.	.	.	86	157
1832.00	*	157	0	.	.	.	86	156
1847.00		156	1	0.7017	0.2983	0.0272	87	155
1874.00	*	155	0	.	.	.	87	154

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
1882.00	*	.	0	.	.	.	87	153
1882.00	*	154	0	.	.	.	87	152
1908.00	*	152	0	.	.	.	87	151
1925.00		151	1	0.6971	0.3029	0.0274	88	150
1932.00	*	150	0	.	.	.	88	149
1945.00	*	149	0	.	.	.	88	148
1951.00	*	148	0	.	.	.	88	147
1967.00	*	147	0	.	.	.	88	146
1978.00	*	146	0	.	.	.	88	145
1979.00	*	145	0	.	.	.	88	144
2022.00	*	144	0	.	.	.	88	143
2033.00	*	143	0	.	.	.	88	142
2050.00	*	142	0	.	.	.	88	141
2055.00		141	1	0.6921	0.3079	0.0276	89	140
2081.00		140	1	0.6872	0.3128	0.0279	90	139
2090.00		139	1	0.6823	0.3177	0.0281	91	138
2105.00		138	1	0.6773	0.3227	0.0283	92	137
2106.00	*	137	0	.	.	.	92	136
2157.00	*	136	0	.	.	.	92	135
2168.00	*	135	0	.	.	.	92	134
2170.00	*	134	0	.	.	.	92	133
2171.00	*	133	0	.	.	.	92	132
2176.00	*	132	0	.	.	.	92	131
2178.00	*	131	0	.	.	.	92	130
2195.00	*	130	0	.	.	.	92	129
2216.00	*	129	0	.	.	.	92	128
2221.00	*	128	0	.	.	.	92	127

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
2224.00		127	1	0.6720	0.3280	0.0286	93	126
2224.00	*	.	0	.	.	.	93	125
2241.00	*	125	0	.	.	.	93	124
2255.00	*	124	0	.	.	.	93	123
2256.00		123	1	0.6665	0.3335	0.0289	94	122
2272.00	*	122	0	.	.	.	94	121
2288.00		121	1	0.6610	0.3390	0.0292	95	120
2294.00	*	120	0	.	.	.	95	119
2297.00		119	1	0.6554	0.3446	0.0295	96	118
2301.00	*	118	0	.	.	.	96	117
2318.00	*	117	0	.	.	.	96	116
2330.00	*	116	0	.	.	.	96	115
2332.00	*	115	0	.	.	.	96	114
2350.00	*	114	0	.	.	.	96	113
2357.00	*	113	0	.	.	.	96	112
2363.00	*	112	0	.	.	.	96	111
2365.00	*	111	0	.	.	.	96	110
2386.00		110	1	0.6495	0.3505	0.0298	97	109
2400.00		109	1	0.6435	0.3565	0.0301	98	108
2419.00		108	1	0.6376	0.3624	0.0304	99	107
2443.00	*	107	0	.	.	.	99	106
2449.00	*	106	0	.	.	.	99	105
2452.00	*	105	0	.	.	.	99	104
2456.00	*	104	0	.	.	.	99	103
2466.00		103	1	0.6314	0.3686	0.0307	100	102
2468.00	*	102	0	.	.	.	100	101
2475.00	*	101	0	.	.	.	100	100

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
2503.00		100	1	0.6251	0.3749	0.0311	101	99
2504.00	*	.	0	.	.	.	101	98
2504.00	*	99	0	.	.	.	101	97
2527.00	*	97	0	.	.	.	101	96
2540.00		96	1	0.6186	0.3814	0.0314	102	95
2555.00	*	95	0	.	.	.	102	94
2556.00	*	94	0	.	.	.	102	93
2563.00	*	93	0	.	.	.	102	92
2573.00	*	92	0	.	.	.	102	91
2574.00	*	91	0	.	.	.	102	90
2576.00	*	90	0	.	.	.	102	89
2580.00	*	89	0	.	.	.	102	88
2583.00		88	1	0.6115	0.3885	0.0318	103	87
2598.00		87	1	0.6045	0.3955	0.0322	104	86
2609.00	*	86	0	.	.	.	104	85
2615.00	*	85	0	.	.	.	104	84
2624.00	*	84	0	.	.	.	104	83
2644.00	*	83	0	.	.	.	104	82
2657.00	*	82	0	.	.	.	104	81
2666.00	*	81	0	.	.	.	104	80
2689.00		80	1	0.5969	0.4031	0.0327	105	79
2692.00	*	79	0	.	.	.	105	78
2713.00	*	78	0	.	.	.	105	77
2721.00	*	77	0	.	.	.	105	76
2769.00		76	1	0.5891	0.4109	0.0332	106	75
2772.00	*	75	0	.	.	.	106	74
2796.00		74	1	0.5811	0.4189	0.0337	107	73

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
2797.00	*	73	0	.	.	.	107	72
2835.00	*	72	0	.	.	.	107	71
2847.00		71	1	0.5729	0.4271	0.0342	108	70
2863.00	*	70	0	.	.	.	108	69
2870.00	*	69	0	.	.	.	108	68
2891.00	*	68	0	.	.	.	108	67
2944.00	*	67	0	.	.	.	108	66
2976.00	*	66	0	.	.	.	108	65
2990.00	*	65	0	.	.	.	108	64
2995.00	*	64	0	.	.	.	108	63
3050.00	*	63	0	.	.	.	108	62
3059.00	*	62	0	.	.	.	108	61
3069.00	*	61	0	.	.	.	108	60
3086.00		60	1	0.5634	0.4366	0.0349	109	59
3090.00		59	1	0.5538	0.4462	0.0356	110	58
3092.00	*	58	0	.	.	.	110	57
3098.00	*	57	0	.	.	.	110	56
3099.00	*	56	0	.	.	.	110	55
3149.00	*	55	0	.	.	.	110	54
3150.00	*	54	0	.	.	.	110	53
3170.00		53	1	0.5434	0.4566	0.0365	111	52
3222.00		52	1	0.5329	0.4671	0.0372	112	51
3239.00	*	51	0	.	.	.	112	50
3244.00		50	1	0.5223	0.4777	0.0380	113	49
3255.00	*	49	0	.	.	.	113	48
3282.00		48	1	0.5114	0.4886	0.0387	114	47
3297.00	*	47	0	.	.	.	114	46

Product-Limit Survival Estimates

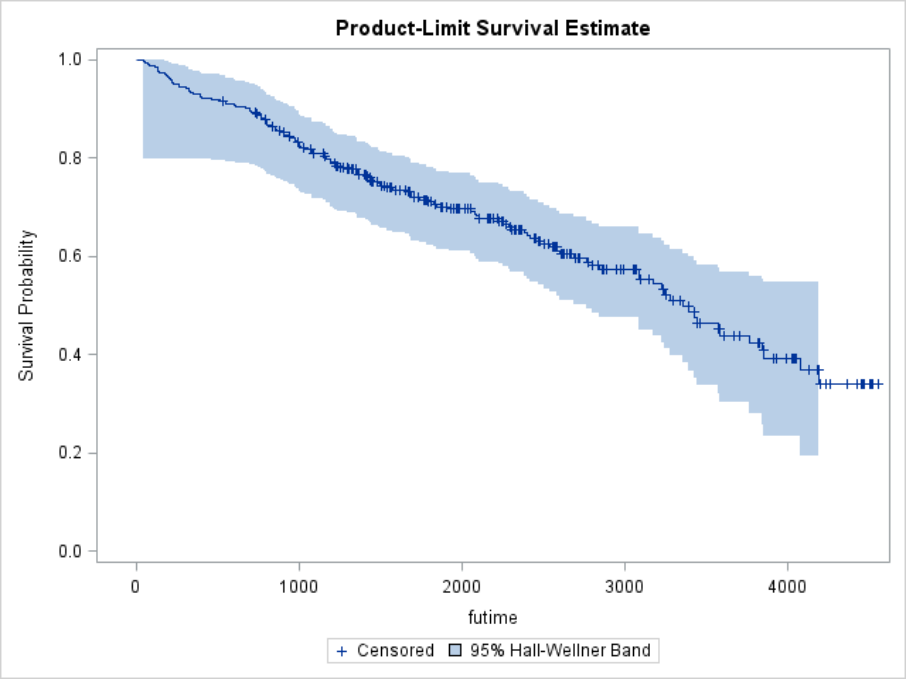
futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
3336.00	*	46	0	.	.	.	114	45
3358.00		45	1	0.5000	0.5000	0.0395	115	44
3388.00	*	44	0	.	.	.	115	43
3395.00		43	1	0.4884	0.5116	0.0402	116	42
3422.00	*	42	0	.	.	.	116	41
3428.00		41	1	0.4765	0.5235	0.0410	117	40
3445.00		40	1	0.4646	0.5354	0.0417	118	39
3445.00	*	.	0	.	.	.	118	38
3458.00	*	38	0	.	.	.	118	37
3574.00		37	1	0.4520	0.5480	0.0424	119	36
3577.00	*	36	0	.	.	.	119	35
3581.00	*	35	0	.	.	.	119	34
3584.00		34	1	0.4387	0.5613	0.0432	120	33
3611.00	*	33	0	.	.	.	120	32
3672.00	*	32	0	.	.	.	120	31
3707.00	*	31	0	.	.	.	120	30
3762.00		30	1	0.4241	0.5759	0.0441	121	29
3820.00	*	29	0	.	.	.	121	28
3823.00	*	28	0	.	.	.	121	27
3839.00		27	1	0.4084	0.5916	0.0452	122	26
3850.00	*	26	0	.	.	.	122	25
3853.00		25	1	0.3921	0.6079	0.0463	123	24
3913.00	*	24	0	.	.	.	123	23
3933.00	*	23	0	.	.	.	123	22
3992.00	*	22	0	.	.	.	123	21
4025.00	*	21	0	.	.	.	123	20
4032.00	*	20	0	.	.	.	123	19

Product-Limit Survival Estimates

futime		Number at Risk	Observed Events	Survival	Failure	Survival Standard Error	Number Failed	Number Left
4039.00	*	19	0	.	.	.	123	18
4050.00	*	18	0	.	.	.	123	17
4079.00		17	1	0.3690	0.6310	0.0490	124	16
4127.00	*	16	0	.	.	.	124	15
4184.00	*	15	0	.	.	.	124	14
4190.00	*	14	0	.	.	.	124	13
4191.00		13	1	0.3406	0.6594	0.0528	125	12
4196.00	*	12	0	.	.	.	125	11
4232.00	*	11	0	.	.	.	125	10
4256.00	*	10	0	.	.	.	125	9
4365.00	*	9	0	.	.	.	125	8
4427.00	*	8	0	.	.	.	125	7
4453.00	*	7	0	.	.	.	125	6
4459.00	*	6	0	.	.	.	125	5
4467.00	*	5	0	.	.	.	125	4
4500.00	*	4	0	.	.	.	125	3
4509.00	*	3	0	.	.	.	125	2
4523.00	*	2	0	.	.	.	125	1
4556.00	*	1	0	0.3406	0.6594	.	125	0

According to above table, we see the table of Kaplan-Meier estimates of the survival function produced by `proc lifetest`. Each row of the table corresponds to an interval of time, beginning at the time in the "FUTIME" column for that row, and ending just before the time in the "`futime`" column in the first subsequent row that has a different "`futime`" value. For example, the time interval represented by the first row is from 0 days to just before 1 day. In this interval, we can see that we had 312 people at risk and that no one died, as "Observed Events" equals 0 and the estimate of the "Survival" function is 1.0000. During the next interval, spanning from 1 day to just before 41 days, 1 people died, indicated by 1 rows of "`futime`" =41.00 and by "Observed Events" =1 in the last row where "`futime`" =41.00. It is important to note that the survival probabilities listed in the Survival column are unconditional and are to be interpreted as the

probability of surviving from the beginning of follow up time up to the number days in the `futime` column. From "`futime`" = 4196 to 4556, we see that there are several records where it appears no events occurred. These are indeed censored observations, further indicated by the "*" appearing in the unlabeled second column. Subjects that are censored after a given time point contribute to the survival function until they drop out of the study but are not counted as a failure. We can see this reflected in the survival function estimate for "`futime`" =1350. During the interval [1350, 1360) 1 out of 214 subjects at-risk died, yielding a conditional probability of survival (the probability of survival in the given interval, given that the subject has survived up to the beginning of the interval) in this interval of (214−1)/214=0.9953. We see that the unconditional probability of surviving beyond 1350 days is 0.7742, since $\hat{S}(1350) = 0.7742$ = p (surviving up to 1350 days) × 0.9953, we can solve for p (surviving up to 1350 days) = 0.7742*0.9953=0.77056. In the table above, we see that the probability surviving beyond 4191 days = 0.3406, the same probability as what we calculated for surviving up to 1350 days, which implies that the censored observations do not change the survival estimates when they leave the study, only the number at risk.



The step function form of the survival function is apparent in the graph of the Kaplan-Meier estimate. When a subject dies at a particular time point, the step function drops, whereas in between failure times the graph remains flat. The survival function drops most steeply at the beginning of study, suggesting that the hazard rate is highest immediately after hospitalization during the first 200 days. Censored observations are represented by vertical ticks on the graph. Notice the survival probability does not change when we encounter a censored observation. Because the observation with the longest follow-up is censored, the survival function will not reach 0. Instead, the survival function will remain at the survival probability estimated at the previous interval. The survival

function is undefined past this final interval at 2358 days. The blue-shaded area around the survival curve represents the 95% confidence band, here Hall-Wellner confidence bands. This confidence band is calculated for the entire survival function, and at any given interval must be wider than the pointwise confidence interval (the confidence interval around a single interval) to ensure that 95% of all pointwise confidence intervals are contained within this band. Many transformations of the survivor function are available for alternate ways of calculating confidence intervals through the `conftype` option, though most transformations should yield very similar confidence intervals.

Survival Function and Cumulative Hazard Rate

			Product-Limit			Nelson-Aalen			
futime	**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
0.00	312	0	1.0000	0	0	0	.	0	312
41.00	312	1	0.9968	0.00321	0.00320	0.00321	0.00321	1	311
51.00	311	1	0.9936	0.00641	0.00452	0.00642	0.00454	2	310
71.00	310	1	0.9904	0.00962	0.00552	0.00965	0.00557	3	309
77.00	309	1	0.9872	0.0128	0.00637	0.0129	0.00644	4	308
110.00	308	1	0.9840	0.0160	0.00711	0.0161	0.00721	5	307
130.00	307	1	0.9808	0.0192	0.00778	0.0194	0.00791	6	306
131.00	306	1	0.9776	0.0224	0.00838	0.0227	0.00856	7	305
140.00	305	1	0.9744	0.0256	0.00895	0.0259	0.00917	8	304
179.00	304	1	0.9712	0.0288	0.00948	0.0292	0.00974	9	303
186.00	303	1	0.9679	0.0321	0.00997	0.0325	0.0103	10	302
191.00	302	1	0.9647	0.0353	0.0104	0.0358	0.0108	11	301
198.00	301	1	0.9615	0.0385	0.0109	0.0392	0.0113	12	300
207.00	300	1	0.9583	0.0417	0.0113	0.0425	0.0118	13	299
216.00	299	1	0.9551	0.0449	0.0117	0.0458	0.0123	14	298
223.00	298	1	0.9519	0.0481	0.0121	0.0492	0.0127	15	297
264.00	.	.	.	.	.	.	.	16	296
264.00	297	2	0.9455	0.0545	0.0129	0.0559	0.0136	17	295
304.00	295	1	0.9423	0.0577	0.0132	0.0593	0.0140	18	294

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
321.00		294	1	0.9391	0.0609	0.0135	0.0627	0.0144	19	293
326.00		293	1	0.9359	0.0641	0.0139	0.0661	0.0148	20	292
334.00		292	1	0.9327	0.0673	0.0142	0.0696	0.0152	21	291
348.00		291	1	0.9295	0.0705	0.0145	0.0730	0.0156	22	290
388.00		290	1	0.9263	0.0737	0.0148	0.0764	0.0159	23	289
400.00		289	1	0.9231	0.0769	0.0151	0.0799	0.0163	24	288
460.00		288	1	0.9199	0.0801	0.0154	0.0834	0.0167	25	287
515.00		287	1	0.9167	0.0833	0.0156	0.0869	0.0170	26	286
533.00	*	286	0	.	.	.	.	.	26	285
549.00		285	1	0.9135	0.0865	0.0159	0.0904	0.0174	27	284
552.00		284	1	0.9102	0.0898	0.0162	0.0939	0.0177	28	283
597.00		283	1	0.9070	0.0930	0.0164	0.0974	0.0181	29	282
611.00		282	1	0.9038	0.0962	0.0167	0.1010	0.0184	30	281
673.00		281	1	0.9006	0.0994	0.0169	0.1045	0.0188	31	280
694.00		280	1	0.8974	0.1026	0.0172	0.1081	0.0191	32	279
708.00		279	1	0.8942	0.1058	0.0174	0.1117	0.0195	33	278
732.00	*	278	0	.	.	.	.	.	33	277
733.00		277	1	0.8909	0.1091	0.0177	0.1153	0.0198	34	276
737.00	*	276	0	.	.	.	.	.	34	275
750.00		275	1	0.8877	0.1123	0.0179	0.1189	0.0201	35	274
762.00		274	1	0.8844	0.1156	0.0181	0.1226	0.0204	36	273
769.00		273	1	0.8812	0.1188	0.0183	0.1262	0.0208	37	272
786.00		272	1	0.8780	0.1220	0.0186	0.1299	0.0211	38	271
788.00	*	271	0	.	.	.	.	.	38	270
790.00		270	1	0.8747	0.1253	0.0188	0.1336	0.0214	39	269

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
797.00		269	1	0.8715	0.1285	0.0190	0.1373	0.0217	40	268
799.00		268	1	0.8682	0.1318	0.0192	0.1411	0.0221	41	267
824.00		267	1	0.8650	0.1350	0.0194	0.1448	0.0224	42	266
837.00	*	266	0	.	.	.	.	.	42	265
839.00	*	265	0	.	.	.	.	.	42	264
850.00		264	1	0.8617	0.1383	0.0196	0.1486	0.0227	43	263
853.00		263	1	0.8584	0.1416	0.0198	0.1524	0.0230	44	262
859.00		262	1	0.8551	0.1449	0.0200	0.1562	0.0233	45	261
877.00	*	261	0	.	.	.	.	.	45	260
890.00		260	1	0.8518	0.1482	0.0202	0.1601	0.0236	46	259
901.00	*	259	0	.	.	.	.	.	46	258
904.00		258	1	0.8485	0.1515	0.0204	0.1639	0.0239	47	257
930.00		257	1	0.8452	0.1548	0.0205	0.1678	0.0243	48	256
939.00	*	256	0	.	.	.	.	.	48	255
943.00		255	1	0.8419	0.1581	0.0207	0.1718	0.0246	49	254
971.00		254	1	0.8386	0.1614	0.0209	0.1757	0.0249	50	253
974.00		253	1	0.8353	0.1647	0.0211	0.1796	0.0252	51	252
980.00		252	1	0.8320	0.1680	0.0213	0.1836	0.0255	52	251
994.00	*	251	0	.	.	.	.	.	52	250
999.00		250	1	0.8287	0.1713	0.0214	0.1876	0.0258	53	249
1000.00		249	1	0.8253	0.1747	0.0216	0.1916	0.0261	54	248
1012.00		248	1	0.8220	0.1780	0.0218	0.1957	0.0264	55	247
1030.00	*	247	0	.	.	.	.	.	55	246
1037.00		246	1	0.8187	0.1813	0.0219	0.1997	0.0268	56	245
1067.00	*	245	0	.	.	.	.	.	56	244

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
1077.00		244	1	0.8153	0.1847	0.0221	0.2038	0.0271	57	243
1080.00		243	1	0.8119	0.1881	0.0223	0.2079	0.0274	58	242
1083.00		242	1	0.8086	0.1914	0.0224	0.2121	0.0277	59	241
1084.00	*	241	0	.	.	.	.	.	59	240
1149.00	*	240	0	.	.	.	.	.	59	239
1152.00		239	1	0.8052	0.1948	0.0226	0.2163	0.0280	60	238
1153.00	*	238	0	.	.	.	.	.	60	237
1165.00		237	1	0.8018	0.1982	0.0227	0.2205	0.0283	61	236
1170.00		236	1	0.7984	0.2016	0.0229	0.2247	0.0286	62	235
1191.00		.	.	.	.	.	.	.	63	234
1191.00		235	2	0.7916	0.2084	0.0232	0.2332	0.0293	64	233
1212.00		233	1	0.7882	0.2118	0.0234	0.2375	0.0296	65	232
1216.00	*	.	0	.	.	.	.	.	65	231
1216.00	*	232	0	.	.	.	.	.	65	230
1217.00		230	1	0.7848	0.2152	0.0235	0.2419	0.0299	66	229
1230.00	*	229	0	.	.	.	.	.	66	228
1234.00	*	228	0	.	.	.	.	.	66	227
1235.00		227	1	0.7813	0.2187	0.0237	0.2463	0.0302	67	226
1250.00	*	226	0	.	.	.	.	.	67	225
1271.00	*	225	0	.	.	.	.	.	67	224
1293.00	*	224	0	.	.	.	.	.	67	223
1295.00	*	223	0	.	.	.	.	.	67	222
1297.00		222	1	0.7778	0.2222	0.0238	0.2508	0.0305	68	221
1300.00	*	221	0	.	.	.	.	.	68	220
1301.00	*	220	0	.	.	.	.	.	68	219

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
1302.00	*	219	0	.	.	.	.	.	68	218
1320.00	*	218	0	.	.	.	.	.	68	217
1321.00	*	217	0	.	.	.	.	.	68	216
1329.00	*	216	0	.	.	.	.	.	68	215
1349.00	*	215	0	.	.	.	.	.	68	214
1350.00		214	1	0.7742	0.2258	0.0240	0.2554	0.0309	69	213
1356.00		213	1	0.7705	0.2295	0.0241	0.2601	0.0313	70	212
1360.00		212	1	0.7669	0.2331	0.0243	0.2649	0.0316	71	211
1363.00	*	.	0	.	.	.	.	.	71	210
1363.00	*	211	0	.	.	.	.	.	71	209
1401.00	*	209	0	.	.	.	.	.	71	208
1408.00	*	208	0	.	.	.	.	.	71	207
1412.00	*	207	0	.	.	.	.	.	71	206
1413.00		206	1	0.7632	0.2368	0.0245	0.2697	0.0320	72	205
1418.00	*	205	0	.	.	.	.	.	72	204
1420.00	*	204	0	.	.	.	.	.	72	203
1427.00		203	1	0.7594	0.2406	0.0246	0.2746	0.0324	73	202
1433.00	*	202	0	.	.	.	.	.	73	201
1434.00		201	1	0.7556	0.2444	0.0248	0.2796	0.0327	74	200
1434.00	*	.	0	.	.	.	.	.	74	199
1435.00	*	199	0	.	.	.	.	.	74	198
1444.00		198	1	0.7518	0.2482	0.0250	0.2847	0.0331	75	197
1447.00	*	197	0	.	.	.	.	.	75	196
1455.00	*	196	0	.	.	.	.	.	75	195
1457.00	*	195	0	.	.	.	.	.	75	194

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
1481.00	*	194	0	.	.	.	.	.	75	193
1487.00		193	1	0.7479	0.2521	0.0251	0.2898	0.0335	76	192
1492.00		192	1	0.7440	0.2560	0.0253	0.2951	0.0339	77	191
1504.00	*	191	0	.	.	.	.	.	77	190
1525.00	*	190	0	.	.	.	.	.	77	189
1536.00		189	1	0.7401	0.2599	0.0255	0.3003	0.0343	78	188
1542.00	*	188	0	.	.	.	.	.	78	187
1558.00	*	187	0	.	.	.	.	.	78	186
1568.00	*	186	0	.	.	.	.	.	78	185
1569.00	*	185	0	.	.	.	.	.	78	184
1576.00		184	1	0.7361	0.2639	0.0256	0.3058	0.0348	79	183
1592.00	*	183	0	.	.	.	.	.	79	182
1614.00	*	.	0	.	.	.	.	.	79	181
1614.00	*	182	0	.	.	.	.	.	79	180
1615.00	*	180	0	.	.	.	.	.	79	179
1656.00	*	179	0	.	.	.	.	.	79	178
1657.00		178	1	0.7319	0.2681	0.0258	0.3114	0.0352	80	177
1666.00	*	177	0	.	.	.	.	.	80	176
1677.00	*	176	0	.	.	.	.	.	80	175
1682.00		175	1	0.7278	0.2722	0.0260	0.3171	0.0357	81	174
1690.00		.	.	.	.	.	.	.	82	173
1690.00		174	2	0.7194	0.2806	0.0264	0.3286	0.0366	83	172
1701.00	*	172	0	.	.	.	.	.	83	171
1702.00	*	171	0	.	.	.	.	.	83	170
1735.00	*	170	0	.	.	.	.	.	83	169

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
1741.00		169	1	0.7151	0.2849	0.0266	0.3345	0.0371	84	168
1765.00	*	168	0	.	.	.	.	.	84	167
1769.00	*	167	0	.	.	.	.	.	84	166
1770.00	*	166	0	.	.	.	.	.	84	165
1776.00	*	165	0	.	.	.	.	.	84	164
1783.00	*	164	0	.	.	.	.	.	84	163
1785.00	*	163	0	.	.	.	.	.	84	162
1786.00		162	1	0.7107	0.2893	0.0268	0.3407	0.0376	85	161
1790.00	*	161	0	.	.	.	.	.	85	160
1810.00	*	160	0	.	.	.	.	.	85	159
1827.00		159	1	0.7063	0.2937	0.0270	0.3470	0.0381	86	158
1831.00	*	158	0	.	.	.	.	.	86	157
1832.00	*	157	0	.	.	.	.	.	86	156
1847.00		156	1	0.7017	0.2983	0.0272	0.3534	0.0386	87	155
1874.00	*	155	0	.	.	.	.	.	87	154
1882.00	*	.	0	.	.	.	.	.	87	153
1882.00	*	154	0	.	.	.	.	.	87	152
1908.00	*	152	0	.	.	.	.	.	87	151
1925.00		151	1	0.6971	0.3029	0.0274	0.3600	0.0392	88	150
1932.00	*	150	0	.	.	.	.	.	88	149
1945.00	*	149	0	.	.	.	.	.	88	148
1951.00	*	148	0	.	.	.	.	.	88	147
1967.00	*	147	0	.	.	.	.	.	88	146
1978.00	*	146	0	.	.	.	.	.	88	145
1979.00	*	145	0	.	.	.	.	.	88	144

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
2022.00	*	144	0	.	.	.	.	.	88	143
2033.00	*	143	0	.	.	.	.	.	88	142
2050.00	*	142	0	.	.	.	.	.	88	141
2055.00		141	1	0.6921	0.3079	0.0276	0.3671	0.0398	89	140
2081.00		140	1	0.6872	0.3128	0.0279	0.3743	0.0405	90	139
2090.00		139	1	0.6823	0.3177	0.0281	0.3814	0.0411	91	138
2105.00		138	1	0.6773	0.3227	0.0283	0.3887	0.0417	92	137
2106.00	*	137	0	.	.	.	.	.	92	136
2157.00	*	136	0	.	.	.	.	.	92	135
2168.00	*	135	0	.	.	.	.	.	92	134
2170.00	*	134	0	.	.	.	.	.	92	133
2171.00	*	133	0	.	.	.	.	.	92	132
2176.00	*	132	0	.	.	.	.	.	92	131
2178.00	*	131	0	.	.	.	.	.	92	130
2195.00	*	130	0	.	.	.	.	.	92	129
2216.00	*	129	0	.	.	.	.	.	92	128
2221.00	*	128	0	.	.	.	.	.	92	127
2224.00		127	1	0.6720	0.3280	0.0286	0.3966	0.0425	93	126
2224.00	*	.	0	.	.	.	.	.	93	125
2241.00	*	125	0	.	.	.	.	.	93	124
2255.00	*	124	0	.	.	.	.	.	93	123
2256.00		123	1	0.6665	0.3335	0.0289	0.4047	0.0432	94	122
2272.00	*	122	0	.	.	.	.	.	94	121
2288.00		121	1	0.6610	0.3390	0.0292	0.4130	0.0440	95	120
2294.00	*	120	0	.	.	.	.	.	95	119

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
2297.00		119	1	0.6554	0.3446	0.0295	0.4214	0.0448	96	118
2301.00	*	118	0	.	.	.	.	.	96	117
2318.00	*	117	0	.	.	.	.	.	96	116
2330.00	*	116	0	.	.	.	.	.	96	115
2332.00	*	115	0	.	.	.	.	.	96	114
2350.00	*	114	0	.	.	.	.	.	96	113
2357.00	*	113	0	.	.	.	.	.	96	112
2363.00	*	112	0	.	.	.	.	.	96	111
2365.00	*	111	0	.	.	.	.	.	96	110
2386.00		110	1	0.6495	0.3505	0.0298	0.4305	0.0457	97	109
2400.00		109	1	0.6435	0.3565	0.0301	0.4396	0.0466	98	108
2419.00		108	1	0.6376	0.3624	0.0304	0.4489	0.0476	99	107
2443.00	*	107	0	.	.	.	.	.	99	106
2449.00	*	106	0	.	.	.	.	.	99	105
2452.00	*	105	0	.	.	.	.	.	99	104
2456.00	*	104	0	.	.	.	.	.	99	103
2466.00		103	1	0.6314	0.3686	0.0307	0.4586	0.0485	100	102
2468.00	*	102	0	.	.	.	.	.	100	101
2475.00	*	101	0	.	.	.	.	.	100	100
2503.00		100	1	0.6251	0.3749	0.0311	0.4686	0.0496	101	99
2504.00	*	.	0	.	.	.	.	.	101	98
2504.00	*	99	0	.	.	.	.	.	101	97
2527.00	*	97	0	.	.	.	.	.	101	96
2540.00		96	1	0.6186	0.3814	0.0314	0.4790	0.0506	102	95
2555.00	*	95	0	.	.	.	.	.	102	94

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
2556.00	*	94	0	.	.	.	.	.	102	93
2563.00	*	93	0	.	.	.	.	.	102	92
2573.00	*	92	0	.	.	.	.	.	102	91
2574.00	*	91	0	.	.	.	.	.	102	90
2576.00	*	90	0	.	.	.	.	.	102	89
2580.00	*	89	0	.	.	.	.	.	102	88
2583.00		88	1	0.6115	0.3885	0.0318	0.4904	0.0519	103	87
2598.00		87	1	0.6045	0.3955	0.0322	0.5019	0.0532	104	86
2609.00	*	86	0	.	.	.	.	.	104	85
2615.00	*	85	0	.	.	.	.	.	104	84
2624.00	*	84	0	.	.	.	.	.	104	83
2644.00	*	83	0	.	.	.	.	.	104	82
2657.00	*	82	0	.	.	.	.	.	104	81
2666.00	*	81	0	.	.	.	.	.	104	80
2689.00		80	1	0.5969	0.4031	0.0327	0.5144	0.0546	105	79
2692.00	*	79	0	.	.	.	.	.	105	78
2713.00	*	78	0	.	.	.	.	.	105	77
2721.00	*	77	0	.	.	.	.	.	105	76
2769.00		76	1	0.5891	0.4109	0.0332	0.5275	0.0562	106	75
2772.00	*	75	0	.	.	.	.	.	106	74
2796.00		74	1	0.5811	0.4189	0.0337	0.5410	0.0578	107	73
2797.00	*	73	0	.	.	.	.	.	107	72
2835.00	*	72	0	.	.	.	.	.	107	71
2847.00		71	1	0.5729	0.4271	0.0342	0.5551	0.0595	108	70
2863.00	*	70	0	.	.	.	.	.	108	69

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
2870.00	*	69	0	.	.	.	.	.	108	68
2891.00	*	68	0	.	.	.	.	.	108	67
2944.00	*	67	0	.	.	.	.	.	108	66
2976.00	*	66	0	.	.	.	.	.	108	65
2990.00	*	65	0	.	.	.	.	.	108	64
2995.00	*	64	0	.	.	.	.	.	108	63
3050.00	*	63	0	.	.	.	.	.	108	62
3059.00	*	62	0	.	.	.	.	.	108	61
3069.00	*	61	0	.	.	.	.	.	108	60
3086.00		60	1	0.5634	0.4366	0.0349	0.5718	0.0618	109	59
3090.00		59	1	0.5538	0.4462	0.0356	0.5887	0.0640	110	58
3092.00	*	58	0	.	.	.	.	.	110	57
3098.00	*	57	0	.	.	.	.	.	110	56
3099.00	*	56	0	.	.	.	.	.	110	55
3149.00	*	55	0	.	.	.	.	.	110	54
3150.00	*	54	0	.	.	.	.	.	110	53
3170.00		53	1	0.5434	0.4566	0.0365	0.6076	0.0668	111	52
3222.00		52	1	0.5329	0.4671	0.0372	0.6268	0.0695	112	51
3239.00	*	51	0	.	.	.	.	.	112	50
3244.00		50	1	0.5223	0.4777	0.0380	0.6468	0.0723	113	49
3255.00	*	49	0	.	.	.	.	.	113	48
3282.00		48	1	0.5114	0.4886	0.0387	0.6677	0.0752	114	47
3297.00	*	47	0	.	.	.	.	.	114	46
3336.00	*	46	0	.	.	.	.	.	114	45
3358.00		45	1	0.5000	0.5000	0.0395	0.6899	0.0785	115	44

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
3388.00	*	44	0	.	.	.	.	.	115	43
3395.00		43	1	0.4884	0.5116	0.0402	0.7132	0.0818	116	42
3422.00	*	42	0	.	.	.	.	.	116	41
3428.00		41	1	0.4765	0.5235	0.0410	0.7375	0.0854	117	40
3445.00		40	1	0.4646	0.5354	0.0417	0.7625	0.0890	118	39
3445.00	*	.	0	.	.	.	.	.	118	38
3458.00	*	38	0	.	.	.	.	.	118	37
3574.00		37	1	0.4520	0.5480	0.0424	0.7896	0.0930	119	36
3577.00	*	36	0	.	.	.	.	.	119	35
3581.00	*	35	0	.	.	.	.	.	119	34
3584.00		34	1	0.4387	0.5613	0.0432	0.8190	0.0975	120	33
3611.00	*	33	0	.	.	.	.	.	120	32
3672.00	*	32	0	.	.	.	.	.	120	31
3707.00	*	31	0	.	.	.	.	.	120	30
3762.00		30	1	0.4241	0.5759	0.0441	0.8523	0.1031	121	29
3820.00	*	29	0	.	.	.	.	.	121	28
3823.00	*	28	0	.	.	.	.	.	121	27
3839.00		27	1	0.4084	0.5916	0.0452	0.8894	0.1095	122	26
3850.00	*	26	0	.	.	.	.	.	122	25
3853.00		25	1	0.3921	0.6079	0.0463	0.9294	0.1166	123	24
3913.00	*	24	0	.	.	.	.	.	123	23
3933.00	*	23	0	.	.	.	.	.	123	22
3992.00	*	22	0	.	.	.	.	.	123	21
4025.00	*	21	0	.	.	.	.	.	123	20
4032.00	*	20	0	.	.	.	.	.	123	19

Survival Function and Cumulative Hazard Rate

				Product-Limit			Nelson-Aalen			
futime		**Number at Risk**	**Observed Events**	**Survival**	**Failure**	**Survival Standard Error**	**Cumulative Hazard**	**Cum Haz Standard Error**	**Number Failed**	**Number Left**
4039.00	*	19	0	.	.	.	.	.	123	18
4050.00	*	18	0	.	.	.	.	.	123	17
4079.00		17	1	0.3690	0.6310	0.0490	0.9882	0.1306	124	16
4127.00	*	16	0	.	.	.	.	.	124	15
4184.00	*	15	0	.	.	.	.	.	124	14
4190.00	*	14	0	.	.	.	.	.	124	13
4191.00		13	1	0.3406	0.6594	0.0528	1.0651	0.1516	125	12
4196.00	*	12	0	.	.	.	.	.	125	11
4232.00	*	11	0	.	.	.	.	.	125	10
4256.00	*	10	0	.	.	.	.	.	125	9
4365.00	*	9	0	.	.	.	.	.	125	8
4427.00	*	8	0	.	.	.	.	.	125	7
4453.00	*	7	0	.	.	.	.	.	125	6
4459.00	*	6	0	.	.	.	.	.	125	5
4467.00	*	5	0	.	.	.	.	.	125	4
4500.00	*	4	0	.	.	.	.	.	125	3
4509.00	*	3	0	.	.	.	.	.	125	2
4523.00	*	2	0	.	.	.	.	.	125	1
4556.00	*	1	0	0.3406	0.6594	.	.	.	125	0

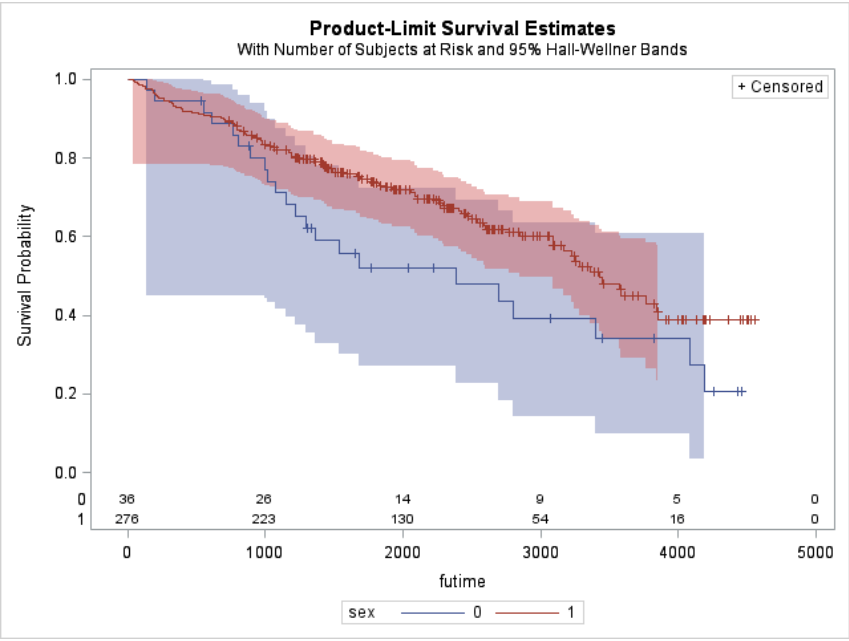
The Nelson-Aalen estimator is requested in SAS through the nelson option on the `proc lifetest` statement. SAS will output both Kaplan Meier estimates of the survival function and Nelson-Aalen estimates of the cumulative hazard function in one table. Researchers are often interested in estimates of survival time at which 50% or 25% of the population have died or failed. Because of the positive skew often seen with follow-up-times, medians are often a better indicator of an "average" survival time. We obtain estimates of these quartiles as well as estimates of the mean survival time by default from proc lifetest. We see that beyond 3395 days, 50% of the

population is expected to have failed. Notice that the interval during which the first 25% of the population is expected to fail, [0,1170) is much shorter than the interval during which the second 25% of the population is expected to fail, [1170,1925). This reinforces our suspicion that the hazard of failure is greater during the beginning of follow-up time.

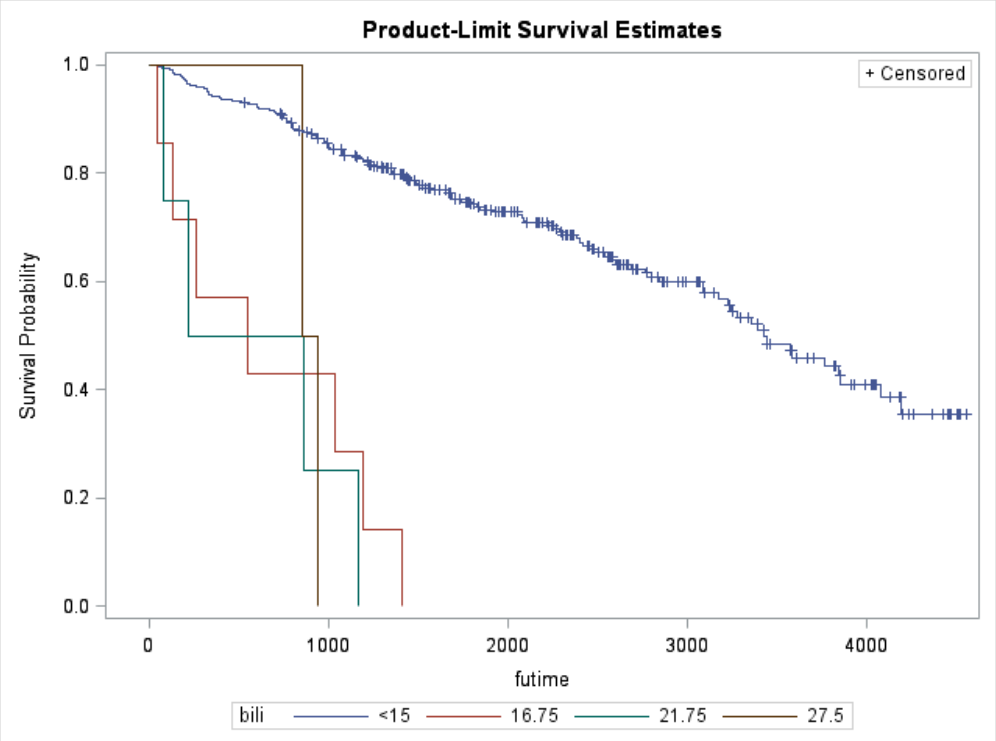
Summary Statistics for Time Variable futime

Quartile Estimates				
Percent	Point Estimate	95% Confidence Interval		
		Transform	[Lower	Upper)
75	.	LOGLOG	4191.00	.
50	3395.00	LOGLOG	3086.00	3839.00
25	1487.00	LOGLOG	1170.00	1925.00

Suppose that you suspect that the survival function is not the same among some of the groups in your study (some groups tend to fail more quickly than others). One can also use non-parametric methods to test for equality of the survival function among groups in the following manner: When provided with a grouping variable in a strata statement in `proc lifetest`, SAS will produce graphs of the survival function (unless other graphs are requested) stratified by the grouping variable as well as tests of equality of the survival function across strata. For example, we could enter the class (categorical) variable gender on the strata statement to request that SAS compare the survival experiences of males and females. In the graph of the Kaplan-Meier estimator stratified by gender below, it appears that females generally have a worse survival experience. This is reinforced by the three significant tests of equality.



We request plots of the hazard function with a bandwidth of 200 days with `plot=hazard (bw=200)`. SAS conveniently allows the creation of strata from a continuous variable, such as bili, on the fly with the strata statement We specify the left endpoints of each bili to form 5 bili categories:15-18.5, 18.5-25, 25-30, 30-40, and >40. The lines in the graph are labeled by the midpoint bili in each group. From the plot we can see that the hazard function indeed appears higher at the beginning of follow-up time and then decreases until it levels off at around 312 days and stays low and mostly constant. The hazard function is also generally higher for the two lowest bili categories. The sudden upticks at the end of follow-up time are not to be trusted, as they are likely due to the few number of subjects at risk at the end. The red curve representing the lowest bili category is truncated on the right because the last person in that group died long before the end of follow-up time.



Next, we are going to fitting a simple Cox regression model for this data. We request Cox regression through `proc phreg` in SAS. Previously, we graphed the survival functions of males in females in our study dataset and suspected that the survival experience after Primary Biliary Cirrhosis may be different between the two genders. Perhaps you also suspect that the hazard rate changes with age as well. Below we demonstrate a simple model in `proc phreg`, where we determine the effects of categorical predictors drug, sex, presence of ascites, presence of hepatomegaly, presence of spiders, presence of edema and continuous predictors, serum bilirubin, serum cholesterol, albumin, urine copper, alkaline phosphatase, SGOT , triglicerides, platelets per

cubic, prothrombin time, histologic stage of disease on the hazard rate. We also would like survival curves based on our model, so we add `plots=survival` to the `proc phreg` statement, although as we shall see this specification is probably insufficient for what we want. On the model statement, on the left side of the equation, we provide the follow up time variable, `futime`, and the censoring variable, `status`, with all censoring values listed in parentheses. On the right side of the equation we list all the predictors.

Model Fit Statistics		
Criterion	**Without Covariates**	**With Covariates**
-2 LOG L	1100.404	931.853
AIC	1100.404	967.853
SBC	1100.404	1016.624

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	168.5510	18	<.0001
Score	295.0615	18	<.0001
Wald	175.8240	18	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
drug	**1**	1	0.17161	0.21919	0.6130	0.4337	1.187	drug 1
age		1	0.02946	0.01163	6.4159	0.0113	1.030	
sex	**0**	1	0.35228	0.30959	1.2948	0.2552	1.422	sex 0
ascites	**0**	1	-0.01492	0.39051	0.0015	0.9695	0.985	ascites 0
hepato	**0**	1	-0.05806	0.25243	0.0529	0.8181	0.944	hepato 0
spiders	**0**	1	-0.06719	0.24878	0.0729	0.7871	0.935	spiders 0
edema	**0**	1	-1.15034	0.41561	7.6608	0.0056	0.317	edema 0
edema	**0.5**	1	-0.89570	0.44874	3.9841	0.0459	0.408	edema 0.5
bili		1	0.08026	0.02619	9.3892	0.0022	1.084	
chol		1	0.0004901	0.0004461	1.2069	0.2719	1.000	
albumin		1	-0.75003	0.30613	6.0026	0.0143	0.472	
copper		1	0.00240	0.00116	4.2510	0.0392	1.002	

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
alk_phos	1	3.12933E-6	0.0000395	0.0063	0.9369	1.000	
sgot	1	0.00386	0.00198	3.7996	0.0513	1.004	
trig	1	-0.0005576	0.00143	0.1524	0.6962	0.999	
platelet	1	0.0007971	0.00119	0.4474	0.5036	1.001	
protime	1	0.24125	0.10680	5.1029	0.0239	1.273	
stage	1	0.44957	0.17468	6.6237	0.0101	1.568	

The above output is only a portion of what SAS produces each time we run `proc phreg`. In particular we would like to highlight the previous tables: Displays fit statistics which are typically used for model comparison and selection. This is our first model, so we have no other model to compare with, except that by default SAS will display model fit statistics of a model with no predictors. We see here that adding gender and particularly age (as we will see below) as predictors improves the fit of the model, as all three statistics decrease. According to testing Global Null Hypothesis, Displays test of hypothesis that all coefficients in the model are 0, that is, an overall test of whether the model as a whole can predict changes in the hazard rate. These tests are asymptotically equivalent, but may differ in smaller samples, in which case the likelihood ratio test is generally preferred. Here the tests agree, and it appears that at least one of our regression coefficients is significantly different from 0. Based on the Analysis of Maximum Likelihood Estimates, Displays model coefficients, tests of significance, and exponential coefficient as hazard ratio. Here it appears that although females have a ~35% (Hazard Ratio = 1.422) increase in the hazard rate compared to males, this decrease is not significant. On the other hand, with each year of age the hazard rate increases by 3% (Hazard Ratio = 1.030), a significant change. Our initial suspicion that the hazard rates were different between genders seems to be wrong once we account for age effects (females are generally older in this dataset), but as shall see the effects are more nuanced. Also notice that there is no intercept. In Cox regression, the intercept is absorbed into the baseline hazard function, which is left unspecified.

6. Conclusions and Future works

According to our sas output, we could know that the covariates AIC values of with covariates is smaller than without covariates and also based on the p values of MLE table, serum bilirubin and Presence of edema have too low values, so these are definitely significant variables. So we can improve our model fitting, remove that significant variables and refit the model. However in this study just we are discussing overall survival analysis and also we discussed cox model fitting. Although many others prefer the type of Cox's PH to more than Zhu M. and Fan G. (2006), there are various factors besides the immune system that affects the lives of patients still. In our last

model. This means that we may leave some information about PBC data not to consider and we could adjust by integrating covariates. Additionally, we do not apply the AFT model to the data that defines time based on covariates. This has a similar procedure, a stressful step that leads to economic growth, which is useful in evaluating the time based on covariates and trust. Secondly, there are many things that are not missing in the PBC data and removing all the documents includes lack of data. But we can consider the arguments that can be attributed to the example of an argument inverse probability weighted (AIPW) by Wang and Chen (2001. When you apply what may happen or its price plan, this feature may be incorporated using EM algorithm method. So, we can review Cox's stress and graphic features by using AIPW schemes to incorporate data that contains incorrect data (or missing data) and more than once. By those parametric and semi-parametric ways, we can find which model is the best model to fit a data for this study.

Acknowledgement

It is with immense pleasure that, I prepared this project report first and foremost my sincere thanks and appreciation to STAT654 course (Fall 2017) instructor Dr. Sanjel Deepak, Professor of Statistics, Department of Mathematics and Statistics, Minnesota state university, Mankato, USA for his tremendous assistance and valuable guidance.

Reference:

1. Liao, C.M. and Tseng, S.T. (2006). Optimal Design for Step-Stress Accelerated Degradation Tests. IEEE TRANSACTIONS ON RELIABILITY, VOL. 55, 59-66.
2. 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.
3. John D. and Ross L. (2002) the Statistical Analysis of Failure Time Data, Wiley.
4. Terry M. and Patricia M. (2000) Modeling Survival Data: Extending the Cox Model, Springer.
5. Zhu M. and Fan G. (2006) Variable Selection in Cox's Proportional Hazards Model Using a Parallel Genetic Algorithm.
6. Introduction to SAS. UCLA: Statistical Consulting Group. from https://stats.idre.ucla.edu/sas/modules/sas-learning-moduleintroduction-to-the features-of sas/ (accessed August 22, 2016).
7. Therneau, TM, Grambsch, PM. (2000). Modeling Survival Data: Extending the Cox Model. Springer: New York. Note: This was the primary reference used for this seminar. It contains numerous examples in SAS and R.
8. Hosmer, DW, Lemeshow, S, May S. (2008). Applied Survival Analysis. Wiley: Hoboken.
9. Lin, DY, Wei, LJ, Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika. 80(30). 557-72.
10. Therneau, TM, Grambsch PM, Fleming TR (1990). Martingale-based residuals for survival models. Biometrika. 77(1). 147-60.

Appendix:

```
options nodate nonumber ps=2000 ls=100;
data pbc;
infile 'E:\STAT654p\pbc.dat.txt';
    input id futime status drug age sex ascites hepato spiders edema
            bili chol albumin copper alk_phos sgot trig platelet
            protime stage;
    age = age/365.25;
if _n_ > 312 then delete;
if status=1 then status=0;
run;

PROC PRINT DATA=pbc;
RUN;

proc lifetest data=pbc atrisk outs=outpbc;
time futime*status(0);
run;


proc lifetest data=pbc atrisk plots=survival(cb) outs=outpbc;
time futime*status(0);
run;


proc lifetest data=pbc atrisk nelson;
time futime*status(0);
run;

proc lifetest data=pbc atrisk nelson;
time futime*status(0);
run;


proc lifetest data=pbc atrisk plots=survival(atrisk cb) outs=outpbc;
strata sex;
time futime*status(0);
run;
```

```
proc lifetest data=pbc atrisk plots=hazard(bw=200) outs=outpbc;
strata bili(15,18.5,25,30,40);
time futime*status(0);
run;

proc phreg data = pbc;
class drug sex ascites hepato spiders edema;
model futime*status(0) =drug age sex ascites hepato spiders edema bili chol
albumin copper alk_phos sgot trig platelet protime stage;
run;

proc phreg data=pbc plots=survival;
class drug sex ascites hepato spiders edema;
model futime*status(0) =drug age sex ascites hepato spiders edema bili chol
albumin copper alk_phos sgot trig platelet protime stage;
run;
```