

Survival Analysis

HOMEWORK 3

(1) Parametric/Accelerated Failure Time Models

For this problem, we will consider the UMARU drug relapse study data described in class. This study consisted of two concurrent randomized trials of residential treatment for drug abuse. At Site A, residents were randomized to either a 3-month or 6-month program of health education and relapse prevention, while Site B compared a 6-month to 12-month program based on a communal living setting. Although 628 participants were enrolled, some had missing values for one or more covariates. To make it easier for you to compare models, the datasets created for the purpose of this assignment have included only the $n = 575$ complete observations. The primary outcome was time to drug relapse or treatment program dropout. The variables include the following:

#	Variable	Type	Len	Label
1	id	Num	8	Subject ID#
2	age	Num	8	Age at entry (yrs)
3	beckscore	Num	8	Beck Total Depression Scale
4	ndrugtx	Num	8	Number previous drug treatments
5	treat	Num	8	Treatment length (0=short,1=long)
6	site	Num	8	Treatment Program (0=A,1=B)
7	los	Num	8	Length of stay in treatment
8	time	Num	8	Time to Drug Relapse or Dropout (days)
9	sensor	Num	8	Drug Relapse (1=yes,0=no)
10	heroin	Num	8	Prior heroin use
11	cocaine	Num	8	Prior cocaine use
12	drugtx	Num	8	Any previous drug treatment
13	ivdrug	Num	8	IV drug use (0=never,1=prev/recent)
14	nonwhite	Num	8	Nonwhite Race (1=yes,0=no)

- (a) Fit the following parametric models to the time to drug relapse or program dropout (**time**), with event indicator **sensor** (called **sensor** but is an event indicator indeed with 1: drug relapse 0: no drug relapse). Using an accelerated failure time (AFT) model approach, including the variables indicated below.

Variables: age, nonwhite, treat, site, ivdrug

Models: Exponential, Weibull, Log-logistic,
Log-normal, Generalized Gamma

- (i) Create a table showing the values of the $-2 \log L$, the total number of parameters (including shape and scale for ϵ) and the AIC for each of these models. Which model appears to provide the best fit to the data on the basis of AIC?

[Note: in Stata, use the `estat ic` command after fitting the model to obtain the AIC.]

- (ii) Does the Weibull model provide an improved fit compared to the exponential model? (justify your answer in at least two ways, using test statistics, confidence intervals, AIC, etc).
- (iii) Given that it includes an extra parameter and thus may allow more flexible fit, does the Generalized Gamma model appear to provide a better fit than any of the models which are nested within it? (Include likelihood ratio test statistics to support your answer)
- (b) Summarize the estimated “time ratio” parameter $\phi = e^\beta$ and its corresponding 95% confidence interval for the IV drug use covariate for each of the above models. Write a brief interpretation of the ϕ parameter from one of these models.
- (c) For the Exponential and Weibull models above, calculate the log-hazard ratio parameter β for the IV drug use covariate, using the β_{AFT} parameter estimates you obtained in (a). Then use these values of β to calculate the estimated HRs for the IV drug use and their corresponding 95% CIs.
- (d) Now Fit a **Cox proportional hazards model** to the time to relapse, including the same five covariates (`age`, `nonwhite`, `treat`, `site`, `ivdrug`). Summarize the effects of these covariates on relapse using estimated hazard ratios and 95% confidence intervals. Compare the hazard ratio for the covariate `ivdrug` from the Cox model to the HR obtained from the Exponential and Weibull models in (c). (Note: these will be hazard ratios adjusting for the other covariates in the model).

Covariate	Model					
	Cox PH		Exponential		Weibull	
	HR	95% CI	HR	95% CI	HR	95% CI
IV drug use						
age			—	—	—	—
nonwhite race			—	—	—	—
treatment			—	—	—	—
site			—	—	—	—

(2) Hazard Rates and Survival from Parametric Models

Now we are going to focus on simple models for the drug relapse data including only the `ivdrug` covariate. The number of residents experiencing drug relapse and mean follow-up time are shown below by IV drug status. The AFT parameter estimates of fitting an Exponential, Weibull, and Log-logistic Model using SAS PROC LIFEREG are also listed at the end of this question.

- Using the information provided above on the number of events and follow-up time, calculate the estimated hazards $\hat{\lambda}_0$ and $\hat{\lambda}_1$ for those without and with prior/recent IV drug use, respectively, under the assumption of an exponential model. Show that these are equivalent to the estimated hazards for each IV drug use group based on the AFT model parameter estimates below.
- Use the estimated hazards from part (a) to calculate the predicted probability of no relapse (survival) by 6 months (182 days) under the Exponential model, for each IV drug use group. At what time point would you expect the predicted survival for those with prior IV drug use to be the same as the 6 month survival in those without prior IV drug use?
- Calculate the predicted survival for the Weibull model at 6 months, for each IV drug use group. Show your calculations! (you should be able to do this from the results provided below).
- Calculate the predicted survival for the log-logistic model at 6 months, for each IV drug use group. Show your calculations! (you should be able to do this from the results provided below).

IV Drug status	# Enrolled	# relapses	Mean Follow-up time (days)
No prior/recent use	223	165	277.031
Prior/recent use	352	299	219.097

Parameter	DF	AFT Beta Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
EXPONENTIAL:							
Intercept	1	5.9254	0.0778	5.7728	6.0779	5793.13	<.0001
ivdrug	1	-0.3727	0.0970	-0.5627	-0.1826	14.77	0.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		
Weibull Shape	0	1.0000	0.0000	1.0000	1.0000		
WEIBULL:							
Intercept	1	5.9255	0.0818	5.7652	6.0858	5250.14	<.0001
ivdrug	1	-0.3813	0.1021	-0.5814	-0.1812	13.95	0.0002
Scale	1	1.0504	0.0404	0.9742	1.1326		
Weibull Shape	1	0.9520	0.0366	0.8829	1.0265		

LOG-LOGISTIC:

Intercept	1	5.3830	0.0863	5.2139	5.5521	3893.11	<.0001
ivdrug	1	-0.3389	0.1090	-0.5525	-0.1252	9.67	0.0019
Scale	1	0.7288	0.0284	0.6752	0.7866		