Maryland State Prison Data (Rossi et al (1980) and Allison (1995:4 & 42))

These data consist of a sample of n = 432 Maryland state prison inmates released from prison in 1980 and followed over a year until the first subsequent arrest.

Covariates include:

- age at release ('**age**' in years)
- employment status (dummy '**wexp**' = 1 if the inmate had FT work experience before incarceration)
- financial assistance (dummy 'fin' = 1 if the inmate received financial aid after release and 0 otherwise); randomly assigned with equal numbers in each category
- marital status (dummy '**mar**' = 1 if inmate was married at the time of release)
- parole status (dummy '**paro**' = 1 if the inmate was released on parole)
- previous conviction status ('**prio**' = number of previous convictions)
- race (dummy '**race**' = 1 if black and 0 otherwise)

In addition, the variable '**week**' reflects the week of the first arrest (if applicable) and the variable '**arrest**' is equal to 1 if 'week' is uncensored and equal to 0 if it is censored.

Fitting the Accelerated Failure Time (AFT) model with the Log-Logistic distribution, we obtain the following.

```
proc lifereg data=one;
  model week*arrest(0)=fin age race wexp mar
        paro prio / dist= dist=llogistic;
run;
```

The LIFEREG Procedure

Model Information

Data Set	WORK.ONE
Dependent Variable	Log(week)
Censoring Variable	arrest
Censoring Value(s)	0
Number of Observations	432
Noncensored Values	114
Right Censored Values	318
Left Censored Values	0
Interval Censored Values	0
Name of Distribution	LLogistic
Log Likelihood	-319.3983709
Number of Observations Read	432
Number of Observations Used	432

Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	3.9183	0.4274	3.0805	4.7561	84.03	<.0001
fin	1	0.2889	0.1456	0.0035	0.5742	3.94	0.0472
age	1	0.0364	0.0156	0.0058	0.0669	5.45	0.0195
race	1	-0.2791	0.2297	-0.7293	0.1710	1.48	0.2242
wexp	1	0.1784	0.1572	-0.1297	0.4865	1.29	0.2563
mar	1	0.3473	0.2697	-0.1812	0.8758	1.66	0.1978

paro	1	0.0508	0.1496	-0.2424	0.3440	0.12	0.7341
prio	1	-0.0692	0.0227	-0.1138	-0.0246	9.25	0.0023
Scale	1	0.6471	0.0559	0.5463	0.7666		

To interpret the parameter estimate for the 'fin' variable, since exp{0.2889} = 1.33, after controlling for the other variables, the expected time to arrest for those who received financial assistance is 1.33 times the expected time to arrest for those who did not received financial assistance.

We should be careful with the above analysis as there may be (marginally) significant interaction between the financial assistance dummy variable and age; this means that the financial incentive program may impact younger and older inmates differently.

```
proc lifereg data=one;
  model week*arrest(0)=fin age race wexp mar
      paro prio fin*age / dist=llogistic;
run;
```

Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	2.5159	1.4527	-0.3314	5.3631	3.00	0.0833
fin	1	-1.0222	0.6752	-2.3456	0.3012	2.29	0.1301
age	1	0.1655	0.1083	-0.0468	0.3777	2.33	0.1265
race	1	-0.2928	0.2298	-0.7432	0.1577	1.62	0.2027
wexp	1	0.1498	0.1586	-0.1610	0.4607	0.89	0.3447
mar	1	0.3034	0.2720	-0.2297	0.8366	1.24	0.2646
paro	1	0.0600	0.1498	-0.2336	0.3536	0.16	0.6887
prio	1	-0.0730	0.0227	-0.1176	-0.0284	10.31	0.0013
fin*age	1	0.0570	0.0290	0.0002	0.1138	3.88	0.0490
Scale	1	0.6435	0.0555	0.5433	0.7621		

Should we be interested in using the semi-parametric Cox Proportional Hazards model, we can use the following program and output.

```
proc phreg data=two;
   model week*arrest(0)=fin age race wexp
         mar paro prio / ties=efron;
run;
```

The PHREG Procedure

Model Information

Data Set	WORK.TWO
Dependent Variable	week
Censoring Variable	arrest
Censoring Value(s)	0
Ties Handling	EFRON
Number of Observations Read	432
Number of Observations Used	432

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
432	114	318	73.61

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1350.761	1317.495
AIC	1350.761	1331.495
SBC	1350.761	1350.649

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	33.2659	7	<.0001
Score	33.5287	7	<.0001
Wald	32.1192	7	<.0001

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
fin	1	-0.37942	0.19138	3.9304	0.0474	0.684
age	1	-0.05743	0.02200	6.8152	0.0090	0.944
race	1	0.31392	0.30799	1.0389	0.3081	1.369
wexp	1	-0.14981	0.21223	0.4983	0.4803	0.861
mar	1	-0.43372	0.38187	1.2900	0.2560	0.648
paro	1	-0.08486	0.19576	0.1879	0.6646	0.919
prio	1	0.09152	0.02865	10.2067	0.0014	1.096

For these data and this model, the estimate of the hazard ratio associated with the 'fin' variable is exp{-0.37942} = 0.684; this means that the hazard of subsequent arrest for those who received financial assistance is 0.684 times the hazard of subsequent arrest for those who did not receive financial help. But again, we should be careful since 'fin*age' again appears to be significant (see below).

```
proc phreg data=two;
   model week*arrest(0)=fin age race wexp mar
        paro prio finage / ties=efron;
run;
```

The PHREG Procedure

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
fin	1	1.62999	1.02908	2.5088	0.1132	5.104
age	1	-0.02269	0.02596	0.7635	0.3822	0.978
race	1	0.32129	0.30825	1.0864	0.2973	1.379
wexp	1	-0.15865	0.21241	0.5579	0.4551	0.853
mar	1	-0.45151	0.38272	1.3918	0.2381	0.637
paro	1	-0.08384	0.19602	0.1829	0.6689	0.920
prio	1	0.09421	0.02848	10.9418	0.0009	1.099
finage	1	-0.08867	0.04508	3.8682	0.0492	0.915

Quite different from the above survival analysis is the following **logistic analysis**. Here, the response variable is whether or not the inmate was subsequently arrested, and non-significant predictor variables have been removed from the model. Note that age enters this model in a quadratic manner.

```
proc logistic descending data=two;
  model arrest=fin age agesq prio educ fin*age;
run;
```

The LOGISTIC Procedure

Model Information

Data Set	WORK.TWO
Response Variable	arrest
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring
Number of Observations Read	432
Number of Observations Used	432

Response Profile

Ordered Value	arrest	Total Frequency
1	1	114
2	0	318

Probability modeled is arrest=1.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	500.602	475.234
SC	504.670	503.713
-2 Log L	498.602	461.234

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	37.3681	6	<.0001
Score	34.7921	6	<.0001
Wald	30.7313	6	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	4.3304	2.2894	3.5779	0.0586
fin	1	1.6266	1.0674	2.3223	0.1275
age	1	-0.3223	0.1644	3.8429	0.0500
agesq	1	0.00513	0.00289	3.1598	0.0755
prio	1	0.0999	0.0371	7.2368	0.0071
educ	1	-0.2522	0.1538	2.6898	0.1010
fin*age	1	-0.0871	0.0452	3.7073	0.0542

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
agesq	1.005	0.999	1.011
prio	1.105	1.027	1.188
educ	0.777	0.575	1.050