

Homework Assignment #5, Applied Survival Data Analysis
Due Monday February 28, 2022

- 1.** Consider the synthetic clinical trial data of survival times from the last homework:

Subject	Survival Time	Censoring Indicator	Group
1	5	1	C
2	8	1	C
3	7	1	T
4	2	1	T
5	9	0	T

(The censoring time for subject 5 is 9 instead of 8; this won't affect the calculations we did in class.)

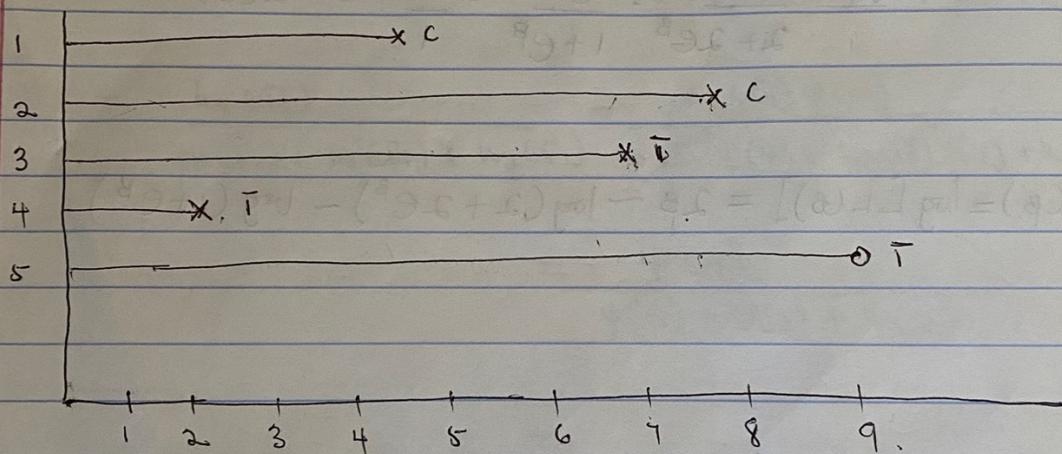
- a. Now write out the partial likelihood for these data, using a proportional hazards model, $h(t; z) = h_0(t)e^{z\beta}$, where $z = 0$ or 1 for the control and treatment groups, respectively. Next, find the log partial likelihood $l(\beta)$. Code the partial likelihood into R as in slides 36, 37, and 38. Plot $l(\beta)$ versus β for a range of values of β from 1 to 2. For example, try this:

```
beta.vec <- seq(-1, 1, by=0.1)
```

Identify the maximum partial likelihood estimate.

Homework Question 1.

subject	SurvTime	death	$z = \text{treatment}$	Covariates	Hazard
1	8	1	0 0	c	$h_0(t) e^{\beta} = h_0(t)$
2	8	1	0 0	c	$h_0(t) e^{\beta} = h_0(t)$
3	7	1	1 1	T	$h_0(t) e^{\beta}$
4	2	1	1	T	$h_0(t) e^{\beta}$
5	9.	0.	1	T	$h_0(t) e^{\beta}$



By ignoring the baseline hazard because they cancel out at both the numerator - and at the denominator:

$$P(1) = \frac{\psi(1)}{\psi(1) + \psi(2) + \psi(3) + \psi(5)} = \frac{1}{1 + 1 + e^{\beta} + e^{\beta}}$$

$$P(1) = \frac{1}{2 + 2e^{\beta}}$$

$$p(2) = \frac{\psi(2)}{\psi(2) + \psi(5)}$$

$$p(2) = \frac{\psi(2)}{\psi(2) + \psi(5)} = \frac{1}{e^B + e^{2B}}$$

$$p(3) = \frac{\psi(3)}{\psi(3) + \psi(2) + \psi(5)} = \frac{e^B}{e^B + 1 + e^B} = \frac{e^B}{2e^B + 1}$$

$$p(4) = \frac{\psi(4)}{\psi(4) + \psi(1) + \psi(2) + \psi(3) + \psi(4) + \psi(5)}$$

$$p(4) = \frac{e^B}{e^B + 1 + 1 + e^B + e^B} = \frac{e^B}{3e^B + 2}$$

$$p(5) = \frac{\psi(5)}{\psi(5)} = \frac{e^B}{e^B}$$

P(5) will be omitted

Because this is a partial likelihood, not a probability,
presumably this term contains little or no information

$$P(5,8,7,2,9) = \frac{\psi(1)}{\psi(1) + \psi(2) + \psi(3) + \psi(5)} \cdot \frac{\psi(2)}{\psi(2) + \psi(5)} \cdot \frac{\psi(3)}{\psi(2) + \psi(3) + \psi(5)}$$

$$\cdot \frac{\psi(4)}{\psi(1) + \psi(2) + \psi(3) + \psi(4) + \psi(5)}$$

PARTIAL LIKELIHOOD.

$$P(5,8,7,2,9) = \frac{1}{2+2e^\beta} \cdot \frac{1}{1+e^\beta} \cdot \frac{e^\beta}{1+2e^\beta} \cdot \frac{e^\beta}{2+3e^\beta}$$

LOG PARTIAL LIKELIHOOD.

$$L(\beta) = \log [L(\beta)] = -2\beta - \log(2+2e^\beta) - \log(1+e^\beta) -$$

$$\log(1+2e^\beta) - \log(2+3e^\beta)$$

Table.

t_i	risk set i c	risk set i T	noi	nu	doi	eli	num	elewm.
2	{1,2}	{3,4,5}	2	3	0	1	e^β	$2+3e^\beta$
5	{1,2}	{3,5}	2	2	1	0	1	$2+2e^\beta$
7	{2}	{3,5}	1	2	0	1	e^β	$1+2e^\beta$
8	{2}	{5}	1	1	1	0	1	$1+e^\beta$
9.	-	{5}	0	1	0	0		censored

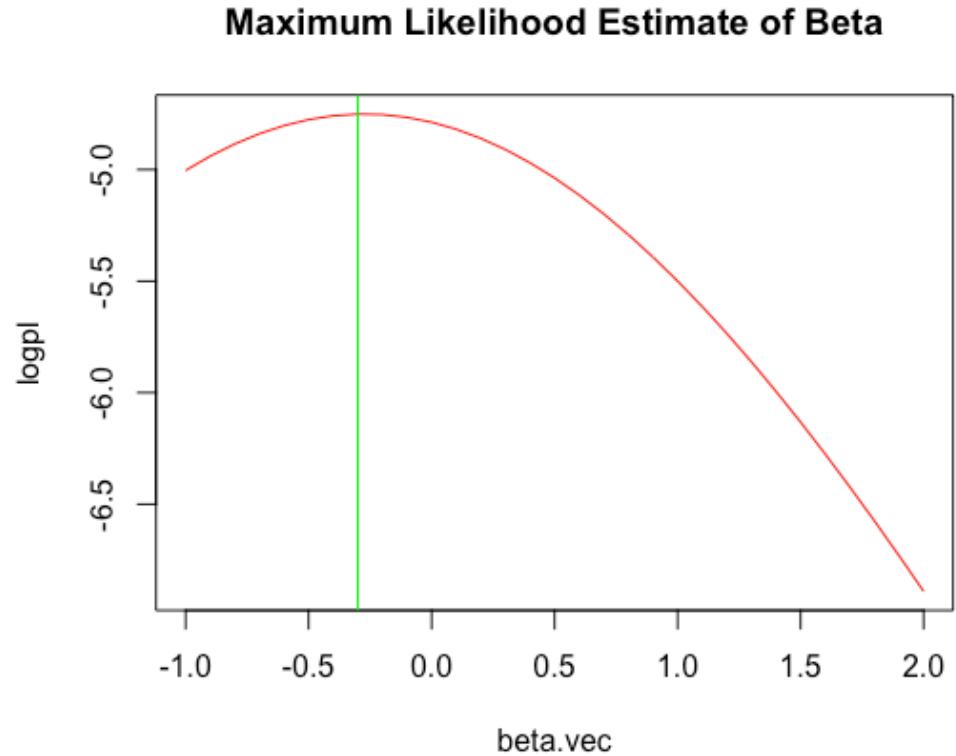
```

logPartialLikelihood <- function(beta) {
  result <- 2*beta - log(2 + 2*exp(beta)) - log(1 + exp(beta)) - log(1 +
  2*exp(beta)) - log(2 + 3*exp(beta))
  result
}

beta.vec <- seq(-1,2, by =0.1)
logpl <- logPartialLikelihood(beta.vec)
plot(logpl ~ beta.vec, type="l", main = "Maximum Likelihood Estimate of
Beta", col = "red")
mle <- beta.vec[which((logPartialLikelihood(beta.vec))==max(logPartialLikelihood(beta.vec)))]
abline(v=mle, col = "green")
mle

result.optim.continious <- optim(par = 1.4, fn = logPartialLikelihood, method
= "BFGS",
control = list(fnscale = -1))
result.optim.continuous

```



Identifying the maximum likelihood estimate using the optim function

```
result.optim.continious <- optim(par = 1.4, fn = logPartialLikelihood, method =  
"BFGS", control = list(fnscale = -1))
```

```
result.optim.continious
```

```
$par
```

```
[1] -0.2739478
```

Therefore, the maximum likelihood estimate is -0.2739478 ≡ -0.3

- b. Fit the proportional hazards model in SAS, and verify that you find the m.p.l.e.:

```
proc phreg;  
  model tt * status(0) = grp ;  
  run;
```

```
data clinical_trial;  
infile "/home/u49706966/AppliedSurvivalAnalysis/Clinical Trial.csv" dsd firstobs=2;  
input tt status group;  
run;
```

```
ODS GRAPHICS On;  
proc phreg;  
  model tt * status(0) = group;  
run;
```

The PHREG Procedure								
Model Information								
Data Set					WORK.CLINICAL_TRIAL			
Dependent Variable					tt			
Censoring Variable					status			
Censoring Value(s)					0			
Ties Handling					BRESLOW			
Number of Observations Read 5 Number of Observations Used 5								
Summary of the Number of Event and Censored Values								
Total	Event	Censored	Percent Censored					
5	4	1	20.00					
Convergence Status								
Convergence criterion (GCONV=1E-8) satisfied.								
Model Fit Statistics								
Criterion	Without Covariates	With Covariates						
-2 LOG L	9.575	9.502						
AIC	9.575	11.502						
SBC	9.575	10.888						
Testing Global Null Hypothesis: BETA=0								
Test	Chi-Square	DF	Pr > ChiSq					
Likelihood Ratio	0.0733	1	0.7866					
Score	0.0739	1	0.7857					
Wald	0.0735	1	0.7863					
Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq			
group	1	-0.27395	1.01067	0.0735	0.7863			
					Hazard Ratio			
					0.760			

Verify that the parameter estimate is the maximum for Part a.

Since the parameter estimates are the same, this verifies that the parameter estimate is the maximum for part a.

c. Compare to the output from fitting the model using the “coxph” function.

```
result.ph <- coxph(Surv(tt, status) ~ grp)
```

```
result.ph
```

```
library(survival)
```

```
#Loading required package: splines
```

```
library(splines)
```

```
tt<- c(5,8,7,2,9)
```

```
death <- c(1,1,1,1,0)
```

```

zz<- c(0,0,1,1,1)

result.ph <- coxph(Surv(tt, death) ~ zz)

summary(result.ph)

```

Call:

coxph(formula = Surv(time, status) ~ group)

n= 5, number of events= 4

coef exp(coef) se(coef) z Pr(>|z|)

group -0.2739 0.7604 1.0107 -0.271 0.786

exp(coef) exp(-coef) lower .95 upper .95

group 0.7604 1.315 0.1049 5.512

Concordance= 0.5 (se = 0.158)

Likelihood ratio test. = 0.07 on 1 df, p=0.8

Wald test = 0.07 on 1 df, p=0.8

Score (logrank) test. = 0.07 on 1 df, p=0.8

d. Show how to get a 95% confidence interval for the hazard ratio, e^β , using the parameter estimate β and its standard error. (Find the 95% confidence interval for β and then exponentiate the endpoints.

library(survival)

#Loading required package: splines

library(splines)

tt<- c(5,8,7,2,9)

death <- c(1,1,1,1,0)

```

zz<- c(0,0,1,1,1)
result.ph <- coxph(Surv(tt, death) ~ zz)
summary(result.ph)

```

Call:

coxph(formula = Surv(tt, death) ~ zz)

n= 5, number of events= 4

<i>coef</i>	<i>exp(coef)</i>	<i>se(coef)</i>	<i>z</i>	<i>Pr(> z)</i>	
zz	-0.2739	0.7604	1.0107	-0.271	0.786

<i>exp(coef)</i>	<i>exp(-coef)</i>	<i>lower .95</i>	<i>upper .95</i>	
zz	0.7604	1.315	0.1049	5.512

Concordance= 0.5 (se = 0.158)

Likelihood ratio test = 0.07 on 1 df, p=0.8

Wald test = 0.07 on 1 df, p=0.8

Score (logrank) test = 0.07 on 1 df, p=0.8

From this

$$e^\beta = 0.7604$$

$$\ln * e^\beta = \ln(0.7604)$$

$$\beta = -0.2739$$

Standard error = 1.0107

*Confidence interval = estimate +/- 1.96 * standard error*

*Confidence interval = -0.2739 +/- 1.96 * 1.0107 = (-2.25487, 1.7072)*

Re-exponentiating the 2 endpoint

$$e^{-2.25487} = 0.1049$$

$$e^{1.7072} = 5.513$$

- 2.** Refer to the “geneConfounder.csv” data set in Canvas (“datasets” folder). Fit a proportional hazards model to the data first with treatment as a predictor, then fit a model with treatment as the predictor and stratify on “genotype”. The code is as follows:

```
result.adj <- coxph(Surv(tt, status) ~ trt)

result.strat <- coxph(Surv(tt, status) ~ trt +
strata(genotype))
```

For each model, examine the parameter estimate of the treatment predictor. Which estimator do you believe? (We discussed this in class.)

```
library(survival)
```

```
setwd("~/Desktop/SPRING 2022/SURVIVAL ANALYSIS/DATA")

geneConfounder <- read.csv("geneConfounder.csv")

result.adj <- coxph(Surv(tt, status)~ trt, data = geneConfounder)

summary(result.adj)
```

Call:

```
coxph(formula = Surv(tt, status) ~ trt, data = geneConfounder)
```

n= 300, number of events= 300

coef exp(coef) se(coef) z Pr(>|z|)

*trt 0.4633 1.5893 0.1173 3.95 7.81e-05 ****

*Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

exp(coef) exp(-coef) lower .95 upper .95

trt 1.589 0.6292 1.263 2

Concordance= 0.572 (se = 0.016)

Likelihood ratio test = 15.47 on 1 df, p=8e-05

Wald test = 15.6 on 1 df, p=8e-05

Score (logrank) test = 15.86 on 1 df, p=7e-05

With treatment as the predictor and stratifying on genotype

strata_genotype <- factor(geneConfounder\$genotype)

result.strata <- coxph(Surv(tt, status) ~ trt + strata(strata_genotype), data = geneConfounder)

result.strata

Call:

coxph(formula = Surv(tt, status) ~ trt + strata(strata_genotype),

data = geneConfounder)

n= 300, number of events= 300

coef exp(coef) se(coef) z Pr(>|z|)

*trt -0.4559 0.6339 0.1645 -2.772 0.00557 ***

*Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

exp(coef) exp(-coef) lower .95 upper .95

trt 0.6339 1.578 0.4592 0.875

Concordance= 0.526 (se = 0.013)

Likelihood ratio test *= 7.74 on 1 df, p=0.005*

Wald test *= 7.68 on 1 df, p=0.006*

Score (logrank) test *= 7.65 on 1 df, p=0.006*

The first test has to do with equality of two groups using the Cox Proportional hazard model. The maximum partial likelihood estimate of beta is 0.4633, the hazard ratio is 1.5893, the standard error is 0.1173, and the p-value is 7.81e-05.

The second test has to do with equality of two groups using the Cox Proportional hazard model. The maximum partial likelihood estimate of beta is -0.4559 , the standard error is 0.1645, hazard ratio of 0.6339 and the p-value is 0.00557.

It can be realized that, both parameters yield results that leads to the rejection of the null hypothesis. however, by checking the coefficients, if you decide not to stratify by genotype, the exp(coefficient) for treatment is greater than 1 and this suggests that the treatment yield a hazard ratio compared to the control. On the other hand, when you evaluate exp(coefficient) value for treatment when you stratified by genotype, it shows an answer less than 1 and this suggests the opposite meaning the treatments yield a lower hazard ratio compared to the control. and this is an example of The Simpsons paradox when you incorporate the confounder into the equation and the result you get is the opposite of what you get without the confounder.

3. For the “geneConfounder.csv” data, fit a Cox model with two covariates:

```
result.trt.gene <- coxph(Surv(tt, status) ~ trt +  
genotype)
```

library(survival)

```
setwd("~/Desktop/SPRING 2022/SURVIVAL ANALYSIS/DATA")
```

```
geneConfounder <- read.csv("geneConfounder.csv")
```

```
result.trt.gen <- coxph(Surv(tt, status) ~ trt + genotype, data = geneConfounder)
```

```
summary(result.trt.gen)
```

Call:

coxph(formula = Surv(tt, status) ~ trt + genotype, data = geneConfounder)

n= 300, number of events= 300

coef exp(coef) se(coef) z Pr(>|z|)

*trt -0.4547 0.6346 0.1635 -2.781 0.00541 ***

*genotypewt -1.5704 0.2080 0.1826 -8.599 < 2e-16 ****

*Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

exp(coef) exp(-coef) lower .95 upper .95

trt 0.6346 1.576 0.4607 0.8743

genotypewt 0.2080 4.809 0.1454 0.2975

Concordance= 0.651 (se = 0.015)

Likelihood ratio test = 93.54 on 2 df, p=<2e-16

Wald test = 91.07 on 2 df, p=<2e-16

Score (logrank) test = 99.07 on 2 df, p=<2e-16

Compare the result to that in Problem 2.

4. Repeat Problems 2 and 3 in SAS, and compare your results.

With treatment as the predictor

proc import out = geneConfounder

datafile = "/home/u49706966/AppliedSurvivalAnalysis/geneConfounder.csv"

DBMS= csv replace;

GETNAMES=Yes;

run;

* 2 coding treatment as predictors;

```
proc phreg data= geneConfounder;
```

```
model tt*status (0) = trt;
```

```
run;
```

The PHREG Procedure								
Model Information								
Data Set		WORK.GENECONFOUNDER						
Dependent Variable		tt						
Censoring Variable		status						
Censoring Value(s)		0						
Ties Handling		BRESLOW						
Number of Observations Read				300				
Number of Observations Used				300				
Summary of the Number of Event and Censored Values								
Total		Event	Censored	Percent Censored				
300		300	0	0.00				
Convergence Status								
Convergence criterion (GCONV=1E-8) satisfied.								
Model Fit Statistics								
Criterion		Without Covariates	With Covariates					
-2 LOG L		2832.476	2817.173					
AIC		2832.476	2819.173					
SBC		2832.476	2822.877					
Testing Global Null Hypothesis: BETA=0								
Test		Chi-Square	DF	Pr > ChiSq				
Likelihood Ratio		15.3030	1	<.0001				
Score		15.6851	1	<.0001				
Wald		15.4315	1	<.0001				
Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq		
trt		1	0.46075	0.11729	15.4315	<.0001		
						1.585		

The PHREG Procedure																															
Model Information																															
Data Set		WORK.GENECONFOUNDER																													
Dependent Variable		tt																													
Censoring Variable		status																													
Censoring Value(s)		0																													
Ties Handling		BRESLOW																													
<table border="1"><tr><td>Number of Observations Read</td><td>300</td></tr><tr><td>Number of Observations Used</td><td>300</td></tr></table>								Number of Observations Read	300	Number of Observations Used	300																				
Number of Observations Read	300																														
Number of Observations Used	300																														
<table border="1"><tr><th colspan="3">Class Level Information</th></tr><tr><th>Class</th><th>Value</th><th>Design Variables</th></tr><tr><td>trt</td><td>0</td><td>0</td></tr><tr><td></td><td>1</td><td>1</td></tr></table>								Class Level Information			Class	Value	Design Variables	trt	0	0		1	1												
Class Level Information																															
Class	Value	Design Variables																													
trt	0	0																													
	1	1																													
<table border="1"><tr><th colspan="4">Summary of the Number of Event and Censored Values</th></tr><tr><th>Total</th><th>Event</th><th>Censored</th><th>Percent Censored</th></tr><tr><td>300</td><td>300</td><td>0</td><td>0.00</td></tr></table>								Summary of the Number of Event and Censored Values				Total	Event	Censored	Percent Censored	300	300	0	0.00												
Summary of the Number of Event and Censored Values																															
Total	Event	Censored	Percent Censored																												
300	300	0	0.00																												
<table border="1"><tr><th colspan="4">Convergence Status</th></tr><tr><td colspan="4">Convergence criterion (GCONV=1E-8) satisfied.</td></tr></table>								Convergence Status				Convergence criterion (GCONV=1E-8) satisfied.																			
Convergence Status																															
Convergence criterion (GCONV=1E-8) satisfied.																															
<table border="1"><tr><th colspan="4">Model Fit Statistics</th></tr><tr><th>Criterion</th><th>Without Covariates</th><th>With Covariates</th><th></th></tr><tr><td>-2 LOG L</td><td>2832.476</td><td>2817.173</td><td></td></tr><tr><td>AIC</td><td>2832.476</td><td>2819.173</td><td></td></tr><tr><td>SBC</td><td>2832.476</td><td>2822.877</td><td></td></tr></table>								Model Fit Statistics				Criterion	Without Covariates	With Covariates		-2 LOG L	2832.476	2817.173		AIC	2832.476	2819.173		SBC	2832.476	2822.877					
Model Fit Statistics																															
Criterion	Without Covariates	With Covariates																													
-2 LOG L	2832.476	2817.173																													
AIC	2832.476	2819.173																													
SBC	2832.476	2822.877																													
<table border="1"><tr><th colspan="4">Testing Global Null Hypothesis: BETA=0</th></tr><tr><th>Test</th><th>Chi-Square</th><th>DF</th><th>Pr > ChiSq</th></tr><tr><td>Likelihood Ratio</td><td>15.3030</td><td>1</td><td><.0001</td></tr><tr><td>Score</td><td>15.6851</td><td>1</td><td><.0001</td></tr><tr><td>Wald</td><td>15.4315</td><td>1</td><td><.0001</td></tr></table>								Testing Global Null Hypothesis: BETA=0				Test	Chi-Square	DF	Pr > ChiSq	Likelihood Ratio	15.3030	1	<.0001	Score	15.6851	1	<.0001	Wald	15.4315	1	<.0001				
Testing Global Null Hypothesis: BETA=0																															
Test	Chi-Square	DF	Pr > ChiSq																												
Likelihood Ratio	15.3030	1	<.0001																												
Score	15.6851	1	<.0001																												
Wald	15.4315	1	<.0001																												
<table border="1"><tr><th colspan="4">Type 3 Tests</th></tr><tr><th>Effect</th><th>DF</th><th>Wald Chi-Square</th><th>Pr > ChiSq</th></tr><tr><td>trt</td><td>1</td><td>15.4315</td><td><.0001</td></tr></table>								Type 3 Tests				Effect	DF	Wald Chi-Square	Pr > ChiSq	trt	1	15.4315	<.0001												
Type 3 Tests																															
Effect	DF	Wald Chi-Square	Pr > ChiSq																												
trt	1	15.4315	<.0001																												
<table border="1"><tr><th colspan="8">Analysis of Maximum Likelihood Estimates</th></tr><tr><th>Parameter</th><th>DF</th><th>Parameter Estimate</th><th>Standard Error</th><th>Chi-Square</th><th>Pr > ChiSq</th><th>Hazard Ratio</th><th>Label</th></tr><tr><td>trt</td><td>1</td><td>0.46075</td><td>0.11729</td><td>15.4315</td><td><.0001</td><td>1.585</td><td>trt 1</td></tr></table>								Analysis of Maximum Likelihood Estimates								Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label	trt	1	0.46075	0.11729	15.4315	<.0001	1.585	trt 1
Analysis of Maximum Likelihood Estimates																															
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label																								
trt	1	0.46075	0.11729	15.4315	<.0001	1.585	trt 1																								

With treatment as the predictor and stratifying on genotype

```
proc phreg data= geneConfounder;
```

```
class trt (ref = "0");
```

```
model tt*status (0) = trt;
```

```
run;
```

The PHREG Procedure

Model Information	
Data Set	WORK.GENECONFOUNDER
Dependent Variable	tt
Censoring Variable	status
Censoring Value(s)	0
Ties Handling	BRESLOW

Number of Observations Read	300
Number of Observations Used	300

Class Level Information		
Class	Value	Design Variables
genotype	mutant	1
	wt	0

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
300	300	0	0.00

Convergence Status			
Convergence criterion (GCONV=1E-8) satisfied.			

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	2832.476	2739.955
AIC	2832.476	2743.955
SBC	2832.476	2751.362

Testing Global Null Hypothesis: BETA=0				
Test	Chi-Square	DF	Pr > ChiSq	
Likelihood Ratio	92.5213	2	<.0001	
Score	97.8863	2	<.0001	
Wald	90.0681	2	<.0001	

Type 3 Tests				
Effect	DF	Wald Chi-Square	Pr > ChiSq	
trt	1	7.6389	0.0057	
genotype	1	73.1639	<.0001	

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
trt		1	-0.45146	0.16334	7.6389	0.0057	0.637	
genotype	mutant	1	1.56093	0.18249	73.1639	<.0001	4.763	genotype mutant

#Fixing a cox model with two covariates

```
proc phreg data= geneConfounder;
class genotype(ref = "mutant");
model tt*status (0) = trt genotype;
```

The PHREG Procedure							
Model Information							
Data Set							WORK.GENECONFOUNDER
Dependent Variable							tt
Censoring Variable							status
Censoring Value(s)							0
Ties Handling							BRESLOW
Number of Observations Read							300
Number of Observations Used							300
Class Level Information							
Class							Design Variables
genotype							0
wt							1
Summary of the Number of Event and Censored Values							
Total							Percent Censored
300							0.00
Convergence Status							
Convergence criterion (GCONV=1E-8) satisfied.							
Model Fit Statistics							
Criterion							Without Covariates
-2 LOG L							2832.476
AIC							2739.955
SBC							2743.955
							2832.476
							2751.362
Testing Global Null Hypothesis: BETA=0							
Test							Chi-Square DF Pr > ChiSq
Likelihood Ratio							92.5213 2 <.0001
Score							97.8863 2 <.0001
Wald							90.0681 2 <.0001
Type 3 Tests							
Effect							DF Wald Chi-Square Pr > ChiSq
trt							1 7.6389 0.0057
genotype							1 73.1639 <.0001
Analysis of Maximum Likelihood Estimates							
Parameter							DF Parameter Estimate Standard Error Chi-Square Pr > ChiSq Hazard Ratio Label
trt							1 -0.45146 0.16334 7.6389 0.0057 0.637
genotype							wt 1 -1.56093 0.18249 73.1639 <.0001 0.210 genotype wt

5. Consider the following modification of the data in problem 1:

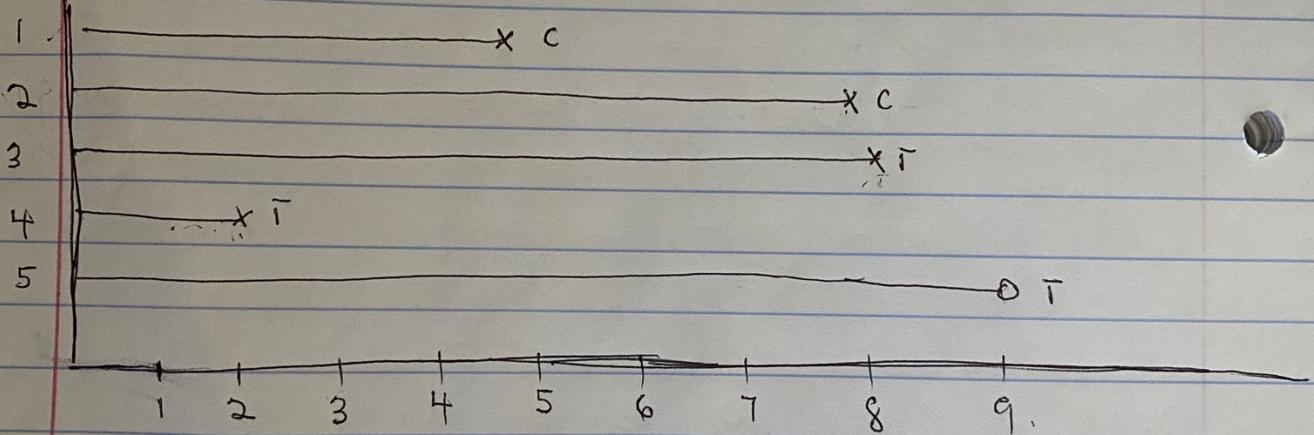
Subject	Survival Time	Censoring Indicator	Group
1	5	1	C
2	8	1	C
3	8	1	T
4	2	1	T
5	9	0	T

Notice that the failure times for Subjects 2 and 3 are now tied. So, which came first? We don't know! If Subject 3 fails slightly BEFORE Subject 2, the partial likelihood will be the same as in problem 1. But if Subject 3 fails slightly AFTER Subject 2, the last factor in the partial likelihood will change. Write down the new partial likelihood with Subject 3 failing slightly after Subject 2.

Homework Question 5.

Index / Subject	SurvTime	death	$\mathbb{I} = \text{treatment group}$	hazard
1	5	1	0	$c h_0(t) e^{\beta} = h_0(t)$
2	8	1	0	$c h_0(t) e^0 = h_0(t)$
3	8	1	1	$\bar{t} h_0(t) e^{\beta}$
4	2	1	1	$\bar{t} h_0(t) e^{\beta}$
5	9	0	1	$\bar{t} h_0(t) e^{\beta}$

Assuming subject 3 failing slightly after subject 2.



We ignore the baseline hazard because they cancel out at both numerator and denominator. The new partial likelihood with 3 failing slightly after subject 2 is given as.

$$P(1) = \frac{\psi(1)}{\psi(1) + \psi(2) + \psi(3) + \psi(5)} = \frac{1}{1 + 1 + e^{\beta} + e^{\beta}}$$

~~$P(1) = \frac{1}{2 + e^{\beta}}$~~

$$P(1) = \frac{1}{2+2e^\beta} \quad (1\bar{1}\bar{1})_{\text{P}} = (P_1, 8, 8, 2)$$

$$P(2) = \frac{\Psi(2)}{\Psi(2)+\Psi(3)+\Psi(5)} \quad (1\bar{1}\bar{1}\bar{1})_{\text{P}} = \frac{1}{1+e^\beta+e^\beta}$$

$$P(2) = \frac{1}{1+2e^\beta} \quad (1\bar{1}\bar{1}\bar{1}\bar{1})_{\text{P}} = (P_2, 8, 8, 2, 2)$$

$$P(3) = \frac{\Psi(3)}{\Psi(3)+\Psi(5)} = \frac{e^\beta}{e^\beta+e^\beta} = 1 \quad (1\bar{1}\bar{1}\bar{1}\bar{1}\bar{1})_{\text{P}} = (P_3, 8, 8, 2, 2, 1)$$

$$P(3) = \frac{e^\beta}{2e^\beta} = \frac{1}{2}$$

$$P(4) = \frac{\Psi(4)}{\Psi(1)+\Psi(2)+\Psi(3)+\Psi(4)+\Psi(5)} = \frac{e^\beta}{1+1+e^\beta+e^\beta+e^\beta}$$

$$P(4) = \frac{e^\beta}{2+3e^\beta}$$

$$P(5) = \frac{\Psi(5)}{\Psi(5)} = \frac{e^\beta}{e^\beta}$$

Because, this is partial likelihood and net probability. Presumably $P(5)$ contain little or no information

$$P(5, 8, 8, 2, 9) = \frac{\psi(1)}{\psi(1) + \psi(2) + \psi(3) + \psi(5)} \cdot \frac{\psi(2)}{\psi(2) + \psi(3) + \psi(5)}$$

$$\frac{\psi(3)}{\psi(3) + \psi(5)} \cdot \frac{\psi(4)}{\psi(1) + \psi(2) + \psi(3) + \psi(4) + \psi(5)}$$

PARTIAL LIKELIHOOD.

$$P(5, 8, 8, 2, 9) = \frac{1}{2+2e^\beta} \cdot \frac{1}{1+2e^\beta} \cdot \frac{e^\beta}{2e^\beta} \cdot \frac{e^\beta}{2+3e^\beta}$$

LOG PARTIAL LIKELIHOOD.

$$\ln(\beta) = \log(1 + e^\beta) = 2\beta - \log(2 + 2e^\beta) - \log(1 + 2e^\beta)$$

$$- \log(2e^\beta) - \log(2 + 3e^\beta)$$

Table.

i	Risk-setic	Risk-setic noi mi doi dili num etenam.	
2	{1, 2}	{3, 4, 5}	2 3 0 1 $(+e^\beta)$ $2+3e^\beta$
5	{1, 2}	{3, 5}	2 2 1 0 1 $(+e^\beta)$ $2+2e^\beta$
8	{2}	{3, 5}	1 2 1 0 1 $(+e^\beta)$ $1+2e^\beta$
8.5	{}	{3, 5}	0 2 0 1 e^β $2e^\beta$
9	{4}	{5}	0 1 0 0 $(+e^\beta)$ Censored.