Stephanie Mecham
2/15/2018

Exercise 6C: Hazard Ratios

Problem 1. Examine the association between blood lead and cardiovascular mortality (**d_cvd**) over time (**pmon_mec**). Use the covariates (confounders) considered above. Check which covariates violate the proportional hazard assumption.

Call:
coxph(formula = Surv(pmon_mec, d_cvd) ~ bpb + age + factor(sex) +
factor(race) + factor(educ) + factor(smk) + factor(alc))

	coef	exp(coef)	se(coef)	z	p
bpb	0.0169	1.0171	0.0118	1.43	0.1521
age	0.1003	1.1055	0.0037	27.13	<2e-16
factor(sex)2	-0.3106	0.7330	0.0960	-3.23	0.0012
factor(race)2	0.2273	1.2552	0.1034	2.20	0.0279
factor(educ)2	0.0720	1.0746	0.0911	0.79	0.4297
factor(educ)3	-0.2390	0.7874	0.1554	-1.54	0.1241
factor(educ)NaN	0.3270	1.3868	0.3843	0.85	0.3948
factor(smk)2	0.0783	1.0814	0.1009	0.78	0.4379
factor(smk)3	0.3997	1.4915	0.1270	3.15	0.0016
factor(alc)1	-0.2994	0.7412	0.1046	-2.86	0.0042
factor(alc)NaN	0.1245	1.1325	0.2118	0.59	0.5568

Likelihood ratio test=1358 on 11 df, p=0
n= 5069, number of events= 576
(5 observations deleted due to missingness)

All factors in the model appeared to have a positive linear relationship with cardiovascular mortality, with the exception of being female, having 16+ years of education, and reporting yes to alcohol consumption. These three factors appear to be protective due to their negative linear relationship with cardiovascular mortality, though it is worth noting that the 16+ years of education factor has a high p-value of 0.1241, which is greater than our alpha level of 0.05 and therefore not likely to be protective at a statistically significant level. Among the factors that appeared to risk factors for cardiovascular mortality, the only factors that were statistically significant (p-value < 0.05) were age, being black, and being a current smoker.

Proportional Hazard Assumption test:

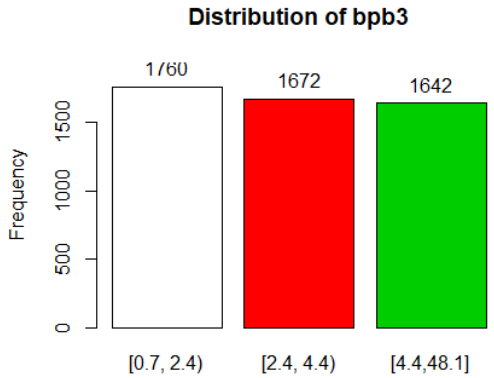
	rho	chisq	p
bpb	-0.02666	0.3798	0.5377
age	0.00832	0.0385	0.8444
factor(sex)2	0.07714	3.4761	0.0623
factor(race)2	-0.01900	0.1936	0.6600
factor(educ)2	0.03268	0.6227	0.4301
factor(educ)3	-0.00957	0.0529	0.8181
factor(educ)NaN	-0.05245	1.5806	0.2087
factor(smk)2	0.04328	1.1095	0.2922
factor(smk)3	0.05015	1.3611	0.2433
factor(alc)1	-0.01541	0.1378	0.7104
factor(alc)NaN	-0.04472	1.1861	0.2761
GLOBAL	NA	10.5836	0.4788

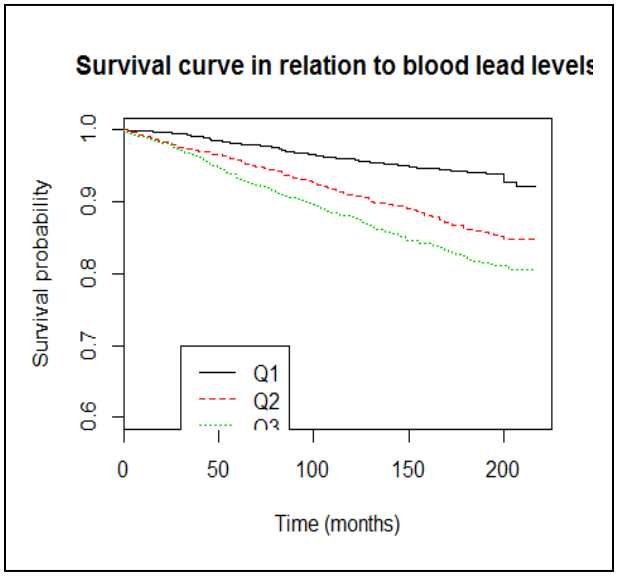
Because all of the p-values are above the alpha threshold of 0.05, this indicates that none of the covariates are in violation of the proportionality assumption.

Problem 2. Compute hazard ratio (HR) (95% CI) per for a 1-SD increase in blood lead using the covariates above.

Hazard Ratio: 1.0562 (0.98, 1.138). We can interpret this to mean that for every 1-SD increase in blood lead levels, rates of cardiovascular mortality events are 5.62% higher in the individuals with the higher blood lead levels. It is relevant to note that this hazard ratio had a p-value of 0.152, which is greater than our standard alpha threshold of 0.05 for determining statistical significance and therefore does not represent a statistically significant hazard ratio.

Problem 3. Compare the lowest vs. the highest tertiles of blood lead using the covariates above. Compute HR and 95% CI.





Comparing the highest tertile of blood lead levels against the lowest tertile resulted in a hazard ratio of 1.2792 (1.0013, 1.6343) We can interpret this to mean that the group in the highest tertile of blood lead levels experienced cardiovascular mortality at rates 0.13% higher compared to the lowest blood level tertile group. It is relevant to note that this hazard ratio had a p-value of 0.04878, which is lower than our standard alpha threshold of 0.05 for determining statistical significance and appears to be (just barely) statistically significant.

Code.

```
>#EPID 674 | HW 6C | Stephanie Mecham
>
>setwd("\\Users\smecham\Desktop\Hw2") #setting working directory
>
>#Installing and Calling Packages
>
>library(foreign)
>library(Hmisc)
>library(stargazer)
>library(sas7bdat)
>install.packages("C:/Users/smecham/Desktop/epicalc_2.15.1.0.tar.gz", repos = NULL, type =
"source")
>install.packages("C:/Users/smecham/Desktop/epicalc_2.
15.1.0.tar.gz", repos = NULL, type = "source") * installing
*source* package 'epicalc' ...
** package 'epicalc' successfully unpacked and MD5 sums
checked
** R
** data
** demo
** preparing package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded
* DONE (epicalc)
>library(epicalc)
> library(survival)
>
>
>#Importing file & attaching
>options(max.print=999999)
>read.sas7bdat("nhanes3.sas7bdat")
> nhanes <- read.sas7bdat("nhanes3.sas7bdat")
> save(nhanes, file="nhanes3.rda") #saving for future use
> attach(nhanes)
>
>#Problem One: Running a Cox Regression with Covariates
> cox.bpb <- coxph(Surv(pmon_mec, d_cvd) ~ bpb + age + factor(sex) + factor(race) +
factor(educ) + factor(smk) + factor(alc))

>coxph(Surv(pmon_mec, d_cvd) ~ bpb + age + factor(sex) + factor(race) + factor(educ) +
factor(smk) + factor(alc))
 > summary(cox.bpb)
 Call:
 coxph(formula = Surv(pmon_mec, d_cvd) ~ bpb + age + factor(sex) +
     factor(race) + factor(educ) + factor(smk) + factor(alc))

   n= 5069, number of events= 576
    (5 observations deleted due to missingness)

                    coef exp(coef)  se(coef)      z Pr(>|z|)
 bpb              0.016945  1.017089  0.011832  1.432  0.15211
 age              0.100316  1.105521  0.003698 27.126  < 2e-16
 factor(sex)2    -0.310570  0.733029  0.096033 -3.234  0.00122
```

```
factor(race)2    0.227329  1.255243  0.103410  2.198  0.02792
factor(educ)2    0.071961  1.074614  0.091118  0.790  0.42967
factor(educ)3   -0.239028  0.787393  0.155431 -1.538  0.12409
factor(educ)NaN  0.327016  1.386824  0.384258  0.851  0.39475
factor(smk)2     0.078291  1.081438  0.100932  0.776  0.43793
factor(smk)3     0.399752  1.491455  0.126951  3.149  0.00164
factor(alc)1    -0.299455  0.741222  0.104547 -2.864  0.00418
factor(alc)NaN   0.124456  1.132532  0.211810  0.588  0.55681

bpb
age             ***
factor(sex)2     **
factor(race)2    *
factor(educ)2
factor(educ)3
factor(educ)NaN
factor(smk)2
factor(smk)3     **
factor(alc)1     **
factor(alc)NaN
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

               exp(coef) exp(-coef) lower .95 upper .95
bpb               1.0171     0.9832    0.9938    1.0410
age               1.1055     0.9046    1.0975    1.1136
factor(sex)2      0.7330     1.3642    0.6073    0.8848
factor(race)2     1.2552     0.7967    1.0250    1.5373
factor(educ)2     1.0746     0.9306    0.8989    1.2847
factor(educ)3     0.7874     1.2700    0.5806    1.0678
factor(educ)NaN   1.3868     0.7211    0.6530    2.9451
factor(smk)2      1.0814     0.9247    0.8873    1.3180
factor(smk)3      1.4915     0.6705    1.1629    1.9128
factor(alc)1      0.7412     1.3491    0.6039    0.9098
factor(alc)NaN    1.1325     0.8830    0.7478    1.7153

Concordance= 0.885  (se = 0.012 )
Rsquare= 0.235   (max possible= 0.849 )
Likelihood ratio test= 1358  on 11 df,   p=0
Wald test            = 878.7  on 11 df,   p=0
Score (logrank) test = 1334  on 11 df,   p=0

>
> #Testing for Proportional Hazards Assumption
> test.prop <- cox.zph(cox.bpb)
> test.prop
                  rho   chisq      p
bpb            -0.02666  0.3798 0.5377
age             0.00832  0.0385 0.8444
factor(sex)2    0.07714  3.4761 0.0623
factor(race)2  -0.01900  0.1936 0.6600
factor(educ)2   0.03268  0.6227 0.4301
factor(educ)3  -0.00957  0.0529 0.8181
factor(educ)NaN -0.05245  1.5806 0.2087
factor(smk)2    0.04328  1.1095 0.2922
factor(smk)3    0.05015  1.3611 0.2433
factor(alc)1   -0.01541  0.1378 0.7104
factor(alc)NaN -0.04472  1.1861 0.2761
```

```
GLOBAL              NA 10.5836 0.4788

>#Standardizing blood level variable so one unit increase = 1-SD unit increase
> zbpb <- (bpb - mean(bpb)) / sd(bpb)
> zcox <- coxph(Surv(pmon_mec, d_cvd) ~ zbpb + age + factor(sex) + factor(race) +
factor(educ) + factor(smk) + factor(alc))
>
>
>#Finding Hazard Ratios & 95% Confidence Intervals
>
>
 > summary(zcox)
 Call:
 coxph(formula = Surv(pmon_mec, d_cvd) ~ zbpb + age + factor(sex) +
     factor(race) + factor(educ) + factor(smk) + factor(alc))

   n= 5069, number of events= 576
    (5 observations deleted due to missingness)

                    coef exp(coef)  se(coef)      z Pr(>|z|)
 zbpb             0.054693  1.056216  0.038190  1.432  0.15211
 age              0.100316  1.105521  0.003698 27.126  < 2e-16
 factor(sex)2    -0.310570  0.733029  0.096033 -3.234  0.00122
 factor(race)2    0.227329  1.255243  0.103410  2.198  0.02792
 factor(educ)2    0.071961  1.074614  0.091118  0.790  0.42967
 factor(educ)3   -0.239028  0.787393  0.155431 -1.538  0.12409
 factor(educ)NaN  0.327016  1.386824  0.384258  0.851  0.39475
 factor(smk)2     0.078291  1.081438  0.100932  0.776  0.43793
 factor(smk)3     0.399752  1.491455  0.126951  3.149  0.00164
 factor(alc)1    -0.299455  0.741222  0.104547 -2.864  0.00418
 factor(alc)NaN   0.124456  1.132532  0.211810  0.588  0.55681

 zbpb
 age             ***
 factor(sex)2     **
 factor(race)2    *
 factor(educ)2
 factor(educ)3
 factor(educ)NaN
 factor(smk)2
 factor(smk)3     **
 factor(alc)1     **
 factor(alc)NaN
 ---
 Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                exp(coef) exp(-coef) lower .95 upper .95
 zbpb              1.0562     0.9468    0.9800    1.1383
 age               1.1055     0.9046    1.0975    1.1136
 factor(sex)2      0.7330     1.3642    0.6073    0.8848
 factor(race)2     1.2552     0.7967    1.0250    1.5373
 factor(educ)2     1.0746     0.9306    0.8989    1.2847
 factor(educ)3     0.7874     1.2700    0.5806    1.0678
 factor(educ)NaN   1.3868     0.7211    0.6530    2.9451
 factor(smk)2      1.0814     0.9247    0.8873    1.3180
 factor(smk)3      1.4915     0.6705    1.1629    1.9128
 factor(alc)1      0.7412     1.3491    0.6039    0.9098
 factor(alc)NaN    1.1325     0.8830    0.7478    1.7153
```

```
Concordance= 0.885  (se = 0.012 )
Rsquare= 0.235   (max possible= 0.849 )
Likelihood ratio test= 1358  on 11 df,   p=0
Wald test            = 878.7  on 11 df,   p=0
Score (logrank) test = 1334  on 11 df,   p=0


>#Problem 3
>#Creating tertiles of blood lead levels
> bpb3<-cut2(bpb, g=3)
 > tab1(bpb3)
 bpb3 :
          Frequency Percent Cum. percent
 [0.7, 2.4)      1760    34.7         34.7
 [2.4, 4.4)      1672    33.0         67.6
 [4.4,48.1]      1642    32.4         100.0
   Total         5074   100.0        100.0
 >

 >#Creating a K-M Life table and curve for blood lead tertiles
>fit.total.bpb3<-survfit(Surv(pmon_mec, d_cvd)~bpb3)
> options(max.print=999999)
> summary(fit.total.bpb3)
>
> plot(fit.total.bpb3, col=c(1:3), lty=c(1:3), mark.time=F, ylim=c(0.6,1.0))
> title(main="Survival curve in relation to blood lead levels", xlab="Time (months)",
ylab="Survival probability")
> legend(30,0.7, legend=c("Q1","Q2","Q3"), lty=c(1:3), col=c(1:3))
>
>#Crude cox analysis with blood lead tertiles
> cox.bpb3<-coxph(Surv(pmon_mec, d_cvd)~bpb3)
> summary(cox.bpb3)
Call:
coxph(formula = Surv(pmon_mec, d_cvd) ~ bpb3)

  n= 5069, number of events= 576
   (5 observations deleted due to missingness)

                coef exp(coef) se(coef)      z Pr(>|z|)
bpb3[2.4, 4.4) 0.8402    2.3168   0.1216 6.909 4.87e-12 ***
bpb3[4.4,48.1] 1.1390    3.1236   0.1173 9.711  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

               exp(coef) exp(-coef) lower .95 upper .95
bpb3[2.4, 4.4)     2.317     0.4316     1.825     2.940
bpb3[4.4,48.1]     3.124     0.3201     2.482     3.931

Concordance= 0.617  (se = 0.012 )
Rsquare= 0.022   (max possible= 0.849 )
Likelihood ratio test= 110.5  on 2 df,   p=0
Wald test            = 94.31  on 2 df,   p=0
Score (logrank) test = 102.7  on 2 df,   p=0
>
>
>#Adjusted cox analysis, taking covariates into account
>
```

```
> cox.bpb3.adj<-coxph(Surv(pmon_mec,d_cvd)~bpb3+age+factor(sex)
+factor(race)+factor(educ)+factor(smk)+factor(alc))
> summary(cox.bpb3.adj)
Call:
coxph(formula = Surv(pmon_mec, d_cvd) ~ bpb3 + age + factor(sex) +
    factor(race) + factor(educ) + factor(smk) + factor(alc))

  n= 5069, number of events= 576
   (5 observations deleted due to missingness)

                    coef exp(coef)  se(coef)      z Pr(>|z|)
bpb3[2.4, 4.4)   0.205342  1.227945  0.123367  1.664  0.09602
bpb3[4.4,48.1]   0.246265  1.279239  0.124973  1.971  0.04878
age              0.099781  1.104929  0.003722 26.809  < 2e-16
factor(sex)2    -0.309090  0.734115  0.095486 -3.237  0.00121
factor(race)2    0.228858  1.257164  0.102869  2.225  0.02610
factor(educ)2    0.072873  1.075594  0.091120  0.800  0.42386
factor(educ)3   -0.233033  0.792127  0.155441 -1.499  0.13383
factor(educ)NaN  0.329578  1.390381  0.384572  0.857  0.39145
factor(smk)2     0.066196  1.068436  0.101045  0.655  0.51240
factor(smk)3     0.388816  1.475233  0.127210  3.056  0.00224
factor(alc)1    -0.304805  0.737267  0.104356 -2.921  0.00349
factor(alc)NaN   0.122506  1.130326  0.211695  0.579  0.56280

bpb3[2.4, 4.4)  .
bpb3[4.4,48.1]  *
age            ***
factor(sex)2    **
factor(race)2   *
factor(educ)2
factor(educ)3
factor(educ)NaN
factor(smk)2
factor(smk)3    **
factor(alc)1    **
factor(alc)NaN
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                exp(coef) exp(-coef) lower .95 upper .95
bpb3[2.4, 4.4)     1.2279     0.8144    0.9642    1.5638
bpb3[4.4,48.1]     1.2792     0.7817    1.0013    1.6343
age                1.1049     0.9050    1.0969    1.1130
factor(sex)2       0.7341     1.3622    0.6088    0.8852
factor(race)2      1.2572     0.7954    1.0276    1.5380
factor(educ)2      1.0756     0.9297    0.8997    1.2859
factor(educ)3      0.7921     1.2624    0.5841    1.0743
factor(educ)NaN    1.3904     0.7192    0.6543    2.9545
factor(smk)2       1.0684     0.9359    0.8765    1.3024
factor(smk)3       1.4752     0.6779    1.1497    1.8930
factor(alc)1       0.7373     1.3564    0.6009    0.9046
factor(alc)NaN     1.1303     0.8847    0.7465    1.7116

Concordance= 0.885  (se = 0.012 )
Rsquare= 0.235   (max possible= 0.849 )
Likelihood ratio test= 1360  on 12 df,   p=0
Wald test            = 877.4  on 12 df,   p=0
Score (logrank) test = 1333  on 12 df,   p=0

> #End of code
```