

Cox PH Models

Landi Luo

12/7/2018

Load packages:

```
if (!require("pacman"))
  install.packages("pacman", repos = "http://cran.us.r-project.org/")
p_load("tidyverse", "survival", "kableExtra")
```

Import data:

```
breast <- readRDS(file = "breast_final.rds")
```

Cox Model: All Covariates

Using the Breslow method of handling ties, we fit a Cox proportional hazards model to the data including all 13 covariates: race, sex, stage, breast subtype, age dx, age, marital status, benign tumor count, malignant tumor count, primary site, pr status, er status, insurance status.

```
fit <- coxph(Surv(SRV_TIME_MON, delta) ~ factor(SEX) + factor(stage) + factor(RAC_RECY) +
  factor(BRST_SUB) + AGE_DX + Age + factor(MAR_STAT) + MALIGCOUNT +
  BENBORDCOUNT + factor(PRIMSITE) + factor(ERSTATUS) + factor(PRSTATUS) +
  factor(INSREC_PUB), data = breast, ties = "breslow" )
summary(fit)
```

```
## Call:
## coxph(formula = Surv(SRV_TIME_MON, delta) ~ factor(SEX) + factor(stage) +
##       factor(RAC_RECY) + factor(BRST_SUB) + AGE_DX + Age + factor(MAR_STAT) +
##       MALIGCOUNT + BENBORDCOUNT + factor(PRIMSITE) + factor(ERSTATUS) +
##       factor(PRSTATUS) + factor(INSREC_PUB), data = breast, ties = "breslow")
##
##      n= 56438, number of events= 3134
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## factor(SEX)2      -0.19493   0.82289   0.21307 -0.915 0.360278
## factor(stage)1       1.83630   6.27327   0.70969  2.587 0.009668 **
## factor(stage)2       3.11544  22.54329   0.70814  4.399 1.09e-05 ***
## factor(stage)3       4.24670  69.87449   0.70810  5.997 2.01e-09 ***
## factor(stage)4       5.90889 368.29651   0.70788  8.347 < 2e-16 ***
## factor(RAC_RECY)2     0.23398   1.26362   0.06608  3.541 0.000399 ***
## factor(RAC_RECY)3     0.05317   1.05461   0.19762  0.269 0.787874
## factor(RAC_RECY)4    -0.23031   0.79429   0.07523 -3.061 0.002203 **
## factor(BRST_SUB)2    -0.96155   0.38230   0.14955 -6.430 1.28e-10 ***
## factor(BRST_SUB)3     0.14835   1.15992   0.06249  2.374 0.017600 *
## factor(BRST_SUB)4    -0.14913   0.86146   0.13985 -1.066 0.286252
## AGE_DX              0.03723   1.03793   0.01366  2.725 0.006429 **
## Age                 -0.01151   0.98856   0.01364 -0.844 0.398708
## factor(MAR_STAT)2    -0.37791   0.68529   0.04874 -7.753 8.94e-15 ***
## factor(MAR_STAT)3    -0.14439   0.86555   0.14404 -1.002 0.316113
## factor(MAR_STAT)4    -0.17069   0.84308   0.06390 -2.671 0.007559 **
## factor(MAR_STAT)5     0.05411   1.05560   0.06282  0.861 0.389114
```

```

## factor(MAR_STAT)6      -0.16544    0.84752    0.33625 -0.492 0.622702
## MALIGCOUNT            0.24579    1.27864    0.05459  4.502 6.72e-06 ***
## BENBORDCOUNT          0.05491    1.05644    0.23297  0.236 0.813686
## factor(PRIMSITE)1      -0.25529    0.77469    0.25544 -0.999 0.317585
## factor(PRIMSITE)2      -0.14997    0.86073    0.25372 -0.591 0.554464
## factor(PRIMSITE)3      -0.04907    0.95212    0.25895 -0.189 0.849712
## factor(PRIMSITE)4      -0.24939    0.77928    0.24770 -1.007 0.314021
## factor(PRIMSITE)5      -0.30684    0.73577    0.25627 -1.197 0.231175
## factor(PRIMSITE)6      -0.23355    0.79171    0.33664 -0.694 0.487824
## factor(PRIMSITE)7      -0.07883    0.92419    0.24796 -0.318 0.750542
## factor(PRIMSITE)8       0.07271    1.07542    0.24772  0.294 0.769135
## factor(ERSTATUS)1       1.05258    2.86504    0.12709  8.282 < 2e-16 ***
## factor(PRSTATUS)1       0.63943    1.89541    0.05220 12.250 < 2e-16 ***
## factor(INSREC_PUB)1    -0.22024    0.80232    0.12278 -1.794 0.072856 .
## factor(INSREC_PUB)2    -0.57835    0.56082    0.12074 -4.790 1.67e-06 ***
## factor(INSREC_PUB)3    -0.43302    0.64855    0.12688 -3.413 0.000643 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## factor(SEX)2          0.8229    1.215223    0.5420    1.2494
## factor(stage)1         6.2733    0.159407    1.5610   25.2104
## factor(stage)2        22.5433    0.044359    5.6266   90.3210
## factor(stage)3        69.8745    0.014311   17.4415  279.9320
## factor(stage)4       368.2965    0.002715   91.9707 1474.8433
## factor(RAC_RECY)2      1.2636    0.791376    1.1101    1.4383
## factor(RAC_RECY)3      1.0546    0.948215    0.7159    1.5535
## factor(RAC_RECY)4      0.7943    1.258990    0.6854    0.9205
## factor(BRST_SUB)2       0.3823    2.615741    0.2852    0.5125
## factor(BRST_SUB)3       1.1599    0.862127    1.0262    1.3111
## factor(BRST_SUB)4       0.8615    1.160826    0.6549    1.1331
## AGE_DX                 1.0379    0.963454    1.0105    1.0661
## Age                   0.9886    1.011577    0.9625    1.0153
## factor(MAR_STAT)2       0.6853    1.459231    0.6229    0.7540
## factor(MAR_STAT)3       0.8655    1.155337    0.6527    1.1479
## factor(MAR_STAT)4       0.8431    1.186125    0.7438    0.9556
## factor(MAR_STAT)5       1.0556    0.947332    0.9333    1.1939
## factor(MAR_STAT)6       0.8475    1.179914    0.4385    1.6382
## MALIGCOUNT            1.2786    0.782083    1.1489    1.4230
## BENBORDCOUNT          1.0564    0.946574    0.6692    1.6678
## factor(PRIMSITE)1       0.7747    1.290836    0.4696    1.2781
## factor(PRIMSITE)2       0.8607    1.161803    0.5235    1.4153
## factor(PRIMSITE)3       0.9521    1.050291    0.5732    1.5816
## factor(PRIMSITE)4       0.7793    1.283239    0.4796    1.2663
## factor(PRIMSITE)5       0.7358    1.359118    0.4453    1.2158
## factor(PRIMSITE)6       0.7917    1.263081    0.4093    1.5315
## factor(PRIMSITE)7       0.9242    1.082023    0.5685    1.5025
## factor(PRIMSITE)8       1.0754    0.929872    0.6618    1.7476
## factor(ERSTATUS)1       2.8650    0.349035    2.2333    3.6755
## factor(PRSTATUS)1       1.8954    0.527591    1.7111    2.0996
## factor(INSREC_PUB)1     0.8023    1.246378    0.6307    1.0206
## factor(INSREC_PUB)2     0.5608    1.783101    0.4426    0.7106
## factor(INSREC_PUB)3     0.6485    1.541907    0.5058    0.8316
##

```

```
## Concordance= 0.892 (se = 0.006 )
## Rsquare= 0.13 (max possible= 0.684 )
## Likelihood ratio test= 7861 on 33 df, p=<2e-16
## Wald test = 7919 on 33 df, p=<2e-16
## Score (logrank) test = 16614 on 33 df, p=<2e-16
```

ANOVA Table: All Covariates

We constructed an Analysis of Variance table to summarize estimates of the risk coefficients and the results of the one degree of freedom tests for each covariate in the model:

```
anova_table <- data.frame(summary(fit)$coefficients)
kable(anova_table, "latex", booktabs = T,
      col.names = c("Coefficient", "Exp. Coeff.", "Std. Error", "Z-Score", "P-Value")) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

	Coefficient	Exp. Coeff.	Std. Error	Z-Score	P-Value
factor(SEX)2	-0.1949277	0.8228942	0.2130742	-0.9148348	0.3602783
factor(stage)1	1.8362975	6.2732683	0.7096869	2.5874756	0.0096682
factor(stage)2	3.1154374	22.5432895	0.7081416	4.3994554	0.0000109
factor(stage)3	4.2467006	69.8744865	0.7080978	5.9973362	0.0000000
factor(stage)4	5.9088884	368.2965141	0.7078797	8.3473060	0.0000000
factor(RAC_REC_Y)2	0.2339824	1.2636223	0.0660793	3.5409354	0.0003987
factor(RAC_REC_Y)3	0.0531742	1.0546133	0.1976203	0.2690724	0.7878740
factor(RAC_REC_Y)4	-0.2303098	0.7942875	0.0752293	-3.0614367	0.0022028
factor(BRST_SUB)2	-0.9615474	0.3823009	0.1495488	-6.4296567	0.0000000
factor(BRST_SUB)3	0.1483530	1.1599223	0.0624927	2.3739237	0.0176002
factor(BRST_SUB)4	-0.1491321	0.8614553	0.1398489	-1.0663806	0.2862516
AGE_DX	0.0372301	1.0379319	0.0136621	2.7250572	0.0064290
Age	-0.0115102	0.9885558	0.0136388	-0.8439307	0.3987081
factor(MAR_STAT)2	-0.3779094	0.6852926	0.0487408	-7.7534431	0.0000000
factor(MAR_STAT)3	-0.1443924	0.8655480	0.1440355	-1.0024780	0.3161128
factor(MAR_STAT)4	-0.1706918	0.8430814	0.0639014	-2.6711726	0.0075587
factor(MAR_STAT)5	0.0541055	1.0555960	0.0628239	0.8612259	0.3891137
factor(MAR_STAT)6	-0.1654417	0.8475193	0.3362467	-0.4920246	0.6227019
MALIGCOUNT	0.2457940	1.2786361	0.0545924	4.5023509	0.0000067
BENBORDCOUNT	0.0549057	1.0564410	0.2329741	0.2356731	0.8136863
factor(PRIMSITE)1	-0.2552902	0.7746917	0.2554352	-0.9994324	0.3175853
factor(PRIMSITE)2	-0.1499728	0.8607314	0.2537248	-0.5910844	0.5544639
factor(PRIMSITE)3	-0.0490668	0.9521175	0.2589473	-0.1894857	0.8497122
factor(PRIMSITE)4	-0.2493870	0.7792784	0.2476974	-1.0068211	0.3140208
factor(PRIMSITE)5	-0.3068361	0.7357712	0.2562656	-1.1973363	0.2311755
factor(PRIMSITE)6	-0.2335542	0.7917147	0.3366432	-0.6937736	0.4878242
factor(PRIMSITE)7	-0.0788329	0.9241944	0.2479612	-0.3179243	0.7505424
factor(PRIMSITE)8	0.0727080	1.0754164	0.2477215	0.2935070	0.7691347
factor(ERSTATUS)1	1.0525836	2.8650436	0.1270924	8.2820327	0.0000000
factor(PRSTATUS)1	0.6394335	1.8954068	0.0521981	12.2501293	0.0000000
factor(INSREC_PUB)1	-0.2202418	0.8023248	0.1227842	-1.7937306	0.0728562
factor(INSREC_PUB)2	-0.5783541	0.5608207	0.1207445	-4.7898992	0.0000017
factor(INSREC_PUB)3	-0.4330201	0.6485475	0.1268777	-3.4128948	0.0006428

Cox Model: Top 9 Significant Variables

Using variable selection methods (LASSO, SCAD, MCP), we decided the top 9 significant variables were:

- Stage
- ERSTATUS
- PRSTATUS
- MALIGCOUNT
- RAC_REC_Y
- PRIMSITE
- BRST_SUB
- MAR_STAT
- AGE_DX

We fit a Cox model with these covariates plus sex:

```
fit2 <- coxph(Surv(SRV_TIME_MON, delta) ~ factor(SEX) + factor(stage) + factor(RAC_REC_Y) +  
             factor(BRST_SUB) + AGE_DX + factor(MAR_STAT) + MALIGCOUNT + factor(PRIMSITE) +  
             factor(ERSTATUS) + factor(PRSTATUS) , data = breast, ties = "breslow" )  
summary(fit2)
```

```
## Call:  
## coxph(formula = Surv(SRV_TIME_MON, delta) ~ factor(SEX) + factor(stage) +  
##       factor(RAC_REC_Y) + factor(BRST_SUB) + AGE_DX + factor(MAR_STAT) +  
##       MALIGCOUNT + factor(PRIMSITE) + factor(ERSTATUS) + factor(PRSTATUS),  
##       data = breast, ties = "breslow")  
##  
##      n= 56438, number of events= 3134  
##  
##              coef exp(coef) se(coef)      z Pr(>|z|)  
## factor(SEX)2      -0.178393  0.836613  0.213092 -0.837 0.402498  
## factor(stage)1      1.826300  6.210864  0.709684  2.573 0.010070 *  
## factor(stage)2      3.119104 22.626102  0.708142  4.405 1.06e-05 ***  
## factor(stage)3      4.273430 71.767360  0.708088  6.035 1.59e-09 ***  
## factor(stage)4      5.947870 382.936859  0.707861  8.403 < 2e-16 ***  
## factor(RAC_REC_Y)2  0.233271  1.262723  0.066092  3.529 0.000416 ***  
## factor(RAC_REC_Y)3  0.092220  1.096606  0.197574  0.467 0.640671  
## factor(RAC_REC_Y)4 -0.210875  0.809875  0.075165 -2.805 0.005024 **  
## factor(BRST_SUB)2 -1.000130  0.367832  0.149293 -6.699 2.10e-11 ***  
## factor(BRST_SUB)3  0.138144  1.148141  0.062471  2.211 0.027013 *  
## factor(BRST_SUB)4 -0.180270  0.835044  0.139628 -1.291 0.196677  
## AGE_DX              0.023703  1.023986  0.001516 15.639 < 2e-16 ***  
## factor(MAR_STAT)2 -0.460123  0.631206  0.047533 -9.680 < 2e-16 ***  
## factor(MAR_STAT)3 -0.137490  0.871543  0.144111 -0.954 0.340057  
## factor(MAR_STAT)4 -0.195014  0.822823  0.063787 -3.057 0.002234 **  
## factor(MAR_STAT)5  0.020216  1.020421  0.062719  0.322 0.747213  
## factor(MAR_STAT)6 -0.218056  0.804081  0.335908 -0.649 0.516239  
## MALIGCOUNT        0.247828  1.281239  0.054547  4.543 5.54e-06 ***  
## factor(PRIMSITE)1 -0.277857  0.757405  0.255384 -1.088 0.276596  
## factor(PRIMSITE)2 -0.181904  0.833682  0.253650 -0.717 0.473286  
## factor(PRIMSITE)3 -0.085534  0.918022  0.258866 -0.330 0.741083  
## factor(PRIMSITE)4 -0.286526  0.750867  0.247603 -1.157 0.247191  
## factor(PRIMSITE)5 -0.344643  0.708473  0.256194 -1.345 0.178547  
## factor(PRIMSITE)6 -0.164242  0.848536  0.335990 -0.489 0.624961  
## factor(PRIMSITE)7 -0.103539  0.901641  0.247887 -0.418 0.676176  
## factor(PRIMSITE)8  0.061991  1.063952  0.247679  0.250 0.802366
```

```

## factor(ERSTATUS)1    1.092302    2.981130    0.126929    8.606    < 2e-16 ***
## factor(PRSTATUS)1    0.634407    1.885904    0.052200    12.153    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## factor(SEX)2          0.8366    1.195295    0.5510    1.2703
## factor(stage)1         6.2109    0.161008    1.5455    24.9595
## factor(stage)2        22.6261    0.044197    5.6473    90.6528
## factor(stage)3        71.7674    0.013934    17.9144   287.5096
## factor(stage)4       382.9369    0.002611    95.6301 1533.4147
## factor(RAC_RECY)2      1.2627    0.791939    1.1093    1.4374
## factor(RAC_RECY)3      1.0966    0.911905    0.7445    1.6152
## factor(RAC_RECY)4      0.8099    1.234758    0.6989    0.9384
## factor(BRST_SUB)2      0.3678    2.718636    0.2745    0.4929
## factor(BRST_SUB)3      1.1481    0.870973    1.0158    1.2977
## factor(BRST_SUB)4      0.8350    1.197541    0.6351    1.0979
## AGE_DX                 1.0240    0.976575    1.0209    1.0270
## factor(MAR_STAT)2      0.6312    1.584270    0.5751    0.6928
## factor(MAR_STAT)3      0.8715    1.147390    0.6571    1.1560
## factor(MAR_STAT)4      0.8228    1.215328    0.7261    0.9324
## factor(MAR_STAT)5      1.0204    0.979987    0.9024    1.1539
## factor(MAR_STAT)6      0.8041    1.243656    0.4163    1.5532
## MALIGCOUNT           1.2812    0.780494    1.1513    1.4258
## factor(PRIMSITE)1      0.7574    1.320298    0.4591    1.2494
## factor(PRIMSITE)2      0.8337    1.199499    0.5071    1.3706
## factor(PRIMSITE)3      0.9180    1.089299    0.5527    1.5248
## factor(PRIMSITE)4      0.7509    1.331793    0.4622    1.2199
## factor(PRIMSITE)5      0.7085    1.411486    0.4288    1.1706
## factor(PRIMSITE)6      0.8485    1.178500    0.4392    1.6393
## factor(PRIMSITE)7      0.9016    1.109089    0.5547    1.4657
## factor(PRIMSITE)8      1.0640    0.939892    0.6548    1.7288
## factor(ERSTATUS)1      2.9811    0.335443    2.3246    3.8232
## factor(PRSTATUS)1      1.8859    0.530250    1.7025    2.0891
##
## Concordance= 0.891 (se = 0.006 )
## Rsquare= 0.129 (max possible= 0.684 )
## Likelihood ratio test= 7788 on 28 df, p=<2e-16
## Wald test = 7848 on 28 df, p=<2e-16
## Score (logrank) test = 16515 on 28 df, p=<2e-16

```