Model Selection

Cindy J. Pang

2025-03-18

```
rm(list=ls())
```

```
source("../src/models.R")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

Cox-PH Models.

```
CoxPH.aic
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ factor(race_cleaned) + factor(FIGO) +
##     age_at_diagnosis, data = data.raceCleaned, ties = "breslow")
##
##                                       coef exp(coef)  se(coef)      z
## factor(race_cleaned)asian           -0.927811  0.395418  0.413392 -2.244
## factor(race_cleaned)black            0.287142  1.332614  0.221872  1.294
## factor(race_cleaned)hispanic         0.086011  1.089818  0.383922  0.224
## factor(race_cleaned)unreported/other -0.906889  0.403778  0.383402 -2.365
## factor(FIGO)Stage II                 0.579678  1.785464  0.660002  0.878
## factor(FIGO)Stage III                1.543566  4.681253  0.580996  2.657
## factor(FIGO)Stage IV                 2.034732  7.650202  0.591420  3.440
## age_at_diagnosis                     0.021626  1.021862  0.004692  4.609
##                                         p
## factor(race_cleaned)asian           0.024808
## factor(race_cleaned)black           0.195603
## factor(race_cleaned)hispanic        0.822733
## factor(race_cleaned)unreported/other 0.018012
## factor(FIGO)Stage II                0.379783
```

```
## factor(FIGO)Stage III                0.007890
## factor(FIGO)Stage IV                 0.000581
## age_at_diagnosis                     4.04e-06
##
## Likelihood ratio test=77.4  on 8 df, p=1.633e-13
## n= 598, number of events= 359
```

```
CoxPH.bic
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ factor(FIGO) + age_at_diagnosis,
##     data = data.raceCleaned, ties = "breslow")
##
##                            coef exp(coef) se(coef)     z        p
## factor(FIGO)Stage II  0.543228  1.721556 0.658470 0.825 0.409380
## factor(FIGO)Stage III 1.533663  4.635122 0.580808 2.641 0.008277
## factor(FIGO)Stage IV  2.050455  7.771436 0.591210 3.468 0.000524
## age_at_diagnosis      0.022861  1.023124 0.004608 4.961 7.01e-07
##
## Likelihood ratio test=61.15  on 4 df, p=1.666e-12
## n= 598, number of events= 359
```

Exponential AFT Models.

```
expAFT.aic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ size.intermediate + factor(race_cleaned) +
##     factor(FIGO) + age_at_diagnosis, data = data.raceCleaned,
##     dist = "exponential")
##
## Coefficients:
##                           (Intercept)                     size.intermediate
##                            11.5786835                            -0.2600141
##            factor(race_cleaned)asian            factor(race_cleaned)black
##                             0.9821891                            -0.3971982
##         factor(race_cleaned)hispanic factor(race_cleaned)unreported/other
##                            -0.1746070                             1.0148055
##                 factor(FIGO)Stage II                factor(FIGO)Stage III
##                            -0.6588743                            -1.7113719
##                 factor(FIGO)Stage IV                      age_at_diagnosis
##                            -2.2627941                            -0.0213537
##
## Scale fixed at 1
##
## Loglik(model)= -3327   Loglik(intercept only)= -3377.4
##  Chisq= 100.91 on 9 degrees of freedom, p= <2e-16
## n= 598
```

```
expAFT.bic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ factor(FIGO) + age_at_diagnosis,
##     data = data.raceCleaned, dist = "exponential")
##
```

```
## Coefficients:
##           (Intercept)  factor(FIGO)Stage II factor(FIGO)Stage III
##           11.52387703           -0.64591491           -1.71559843
##  factor(FIGO)Stage IV      age_at_diagnosis
##           -2.32689756           -0.02338033
##
## Scale fixed at 1
##
## Loglik(model)= -3339.7   Loglik(intercept only)= -3377.4
##  Chisq= 75.54 on 4 degrees of freedom, p= 1.53e-15
## n= 598
```

Weibull AFT Models.

```
weibullAFT.aic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ size.intermediate + factor(race_cleaned) +
##     factor(FIGO) + age_at_diagnosis, data = data.raceCleaned,
##     dist = "weibull")
##
## Coefficients:
##                           (Intercept)                    size.intermediate
##                           12.36301594                           -0.28985830
##            factor(race_cleaned)asian            factor(race_cleaned)black
##                            1.22733827                           -0.44097742
##         factor(race_cleaned)hispanic factor(race_cleaned)unreported/other
##                           -0.20784887                            1.24874844
##                 factor(FIGO)Stage II                factor(FIGO)Stage III
##                           -0.82430720                           -2.11463119
##                 factor(FIGO)Stage IV                     age_at_diagnosis
##                           -2.76756691                           -0.02613169
##
## Scale= 1.294644
##
## Loglik(model)= -3309.1   Loglik(intercept only)= -3352.5
##  Chisq= 86.94 on 9 degrees of freedom, p= 6.68e-15
## n= 598
```

```
weibullAFT.bic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ factor(FIGO) + age_at_diagnosis,
##     data = data.raceCleaned, dist = "weibull")
##
## Coefficients:
##           (Intercept)  factor(FIGO)Stage II factor(FIGO)Stage III
##           12.35580783           -0.80343728           -2.13537950
##  factor(FIGO)Stage IV      age_at_diagnosis
##           -2.86714404           -0.02884478
##
## Scale= 1.311245
##
## Loglik(model)= -3319.9   Loglik(intercept only)= -3352.5
```

```
##  Chisq= 65.22 on 4 degrees of freedom, p= 2.31e-13
## n= 598
```

Log-Logistic AFT Models.

```
loglogisticAFT.aic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ factor(race_cleaned) +
##     factor(FIGO) + age_at_diagnosis, data = data.raceCleaned,
##     dist = "loglogistic")
##
## Coefficients:
##                           (Intercept)            factor(race_cleaned)asian
##                           11.75168203                           1.24102582
##            factor(race_cleaned)black         factor(race_cleaned)hispanic
##                           -0.48585138                           -0.17884472
## factor(race_cleaned)unreported/other                 factor(FIGO)Stage II
##                            1.09502164                           -0.72098231
##                factor(FIGO)Stage III                 factor(FIGO)Stage IV
##                           -1.96859618                           -2.69683380
##                     age_at_diagnosis
##                           -0.03253153
##
## Scale= 0.9524551
##
## Loglik(model)= -3287.3   Loglik(intercept only)= -3330.7
##  Chisq= 86.94 on 8 degrees of freedom, p= 1.94e-15
## n= 598
```

```
loglogisticAFT.bic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ factor(FIGO) + age_at_diagnosis,
##     data = data.raceCleaned, dist = "loglogistic")
##
## Coefficients:
##           (Intercept)  factor(FIGO)Stage II factor(FIGO)Stage III
##           11.93419596           -0.69099201           -1.98625823
##  factor(FIGO)Stage IV      age_at_diagnosis
##           -2.75224858           -0.03446027
##
## Scale= 0.9666854
##
## Loglik(model)= -3296.1   Loglik(intercept only)= -3330.7
##  Chisq= 69.27 on 4 degrees of freedom, p= 3.24e-14
## n= 598
```

Log-Normal AFT Models.

```
lognormalAFT.aic
```

```
## Call:
```

```
## survreg(formula = Surv(time, delta) ~ factor(race_cleaned) +
##     factor(FIGO) + age_at_diagnosis, data = data.raceCleaned,
##     dist = "lognormal")
##
## Coefficients:
##                           (Intercept)            factor(race_cleaned)asian
##                           11.60283189                           1.32670368
##            factor(race_cleaned)black         factor(race_cleaned)hispanic
##                           -0.60372256                          -0.21390418
## factor(race_cleaned)unreported/other                 factor(FIGO)Stage II
##                            0.71779553                          -0.27403185
##                 factor(FIGO)Stage III                 factor(FIGO)Stage IV
##                           -1.53718723                          -2.39908988
##                      age_at_diagnosis
##                           -0.03567528
##
## Scale= 1.698155
##
## Loglik(model)= -3291.4   Loglik(intercept only)= -3333.2
##  Chisq= 83.54 on 8 degrees of freedom, p= 9.43e-15
## n= 598
```

```
lognormalAFT.bic
```

```
## Call:
## survreg(formula = Surv(time, delta) ~ factor(FIGO) + age_at_diagnosis,
##     data = data.raceCleaned, dist = "lognormal")
##
## Coefficients:
##           (Intercept)  factor(FIGO)Stage II factor(FIGO)Stage III
##            11.6248024            -0.1457258            -1.4340866
##  factor(FIGO)Stage IV      age_at_diagnosis
##            -2.3211559            -0.0371437
##
## Scale= 1.716158
##
## Loglik(model)= -3299.2   Loglik(intercept only)= -3333.2
##  Chisq= 68 on 4 degrees of freedom, p= 6.01e-14
## n= 598
```