

Math 150 Survival Analysis Project Test Code

Carrie Saada

May 2, 2019

```
knitr::opts_chunk$set(message=FALSE, warning=FALSE, fig.height=4, fig.width=5,
                        fig.align = "center")
library(tidyverse)
library(broom)
library(survival)
library(survminer)
library(coxed)
```

Importing data:

```
AD = read.csv("AIDSdata.csv")
```

Trying to do power analysis

```
null = sim.survdata(851, 100, x = 1, beta = 0, censor = (851 - 69)/851)
alt = sim.survdata(851, 100, x = 1, beta = -1.454, censor = (851 - 69)/851)
#alt$data$y
cx = coxph(Surv(y, failed) ~ X, data = alt$data)
#coxph(Surv(y, failed) ~ X, data = null$data)
scx = summary(cx)
pval = scx$waldtest[3]
```

X might be the number of covariates?

```
loopVals = c(1:50)
output = rep(NA, length(loopVals))
#output
for(i in loopVals){
  loopSim = sim.survdata(851, 100, x = 1, beta = -1.454, censor = (851 - 69)/851)
  pval = summary(coxph(Surv(y, failed) ~ X, data = loopSim$data))$waldtest[3]
  output[i] = pval < 0.05
}
sum(output) / length(output)
```

```
betaVals = c(0, -0.2, -0.5, -1, -1.454, -1.5, -2, -3)
powersForBeta = rep(NA, length(betaVals))
outerLoop = c(1:length(betaVals)) #loop through betas
innerLoop = c(1:1000) #replicates
for(i in outerLoop){
  pBools = rep(NA, length(innerLoop))
  for(j in innerLoop){
    loopSim = sim.survdata(851, 100, x = 1, beta = betaVals[i], censor = (851 - 69)/851)
```

```

    pval = summary(coxph(Surv(y, failed) ~ X, data = loopSim$data))$waldtest[3]
    pBools[j] = pval < 0.05
  }
  powersForBeta[i] = sum(pBools) / length(pBools)
}

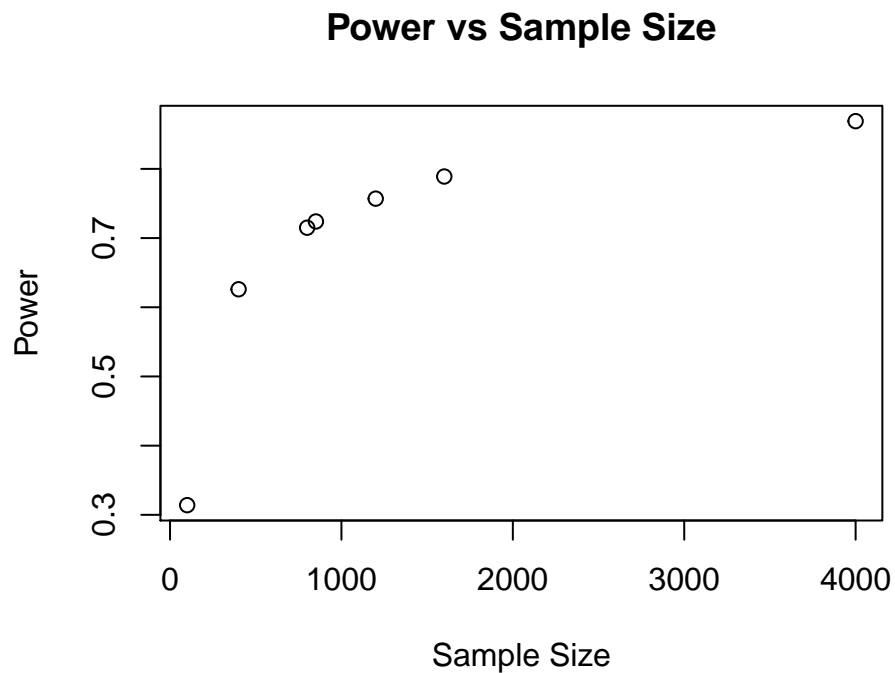
powersForBeta

popVals = c(100, 400, 800, 851, 1200, 1600, 4000)
powersForPop = rep(NA, length(popVals))
outerLoop = c(1:length(popVals)) #loop through betas
innerLoop = c(1:1000) #replicates
for(i in outerLoop){
  pBools = rep(NA, length(innerLoop))
  for(j in innerLoop){
    loopSim = sim.survdata(popVals[i], 100, x= 1, beta = -1.454, censor = (851 - 69)/851)
    pval = summary(coxph(Surv(y, failed) ~ X, data = loopSim$data))$waldtest[3]
    pBools[j] = pval < 0.05
  }
  powersForPop[i] = sum(pBools) / length(pBools)
}

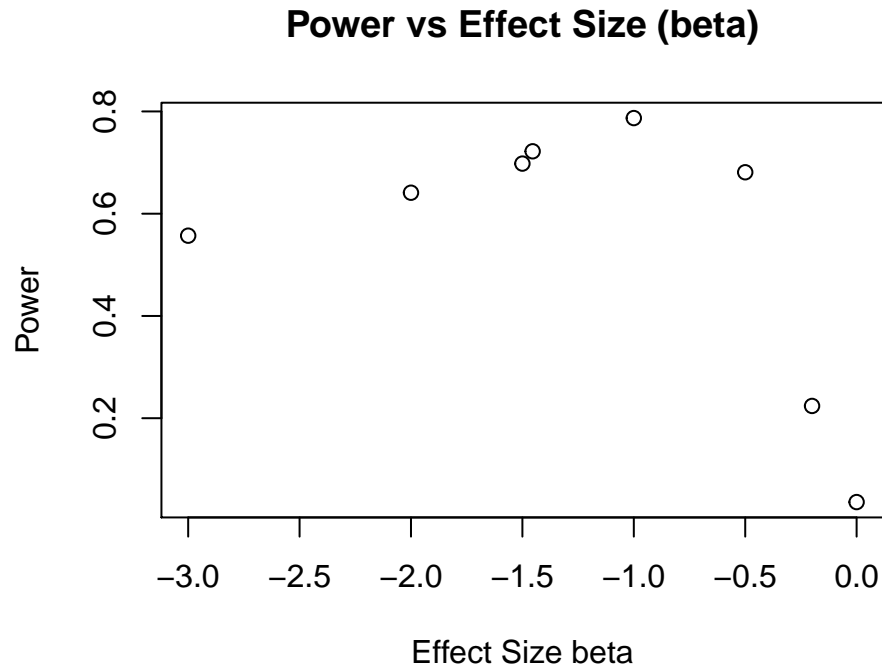
powersForPop

popx = c(100, 400, 800, 851, 1200, 1600, 4000)
popy = c(0.314, 0.626, 0.715, 0.724, 0.757, 0.789, 0.869)
plot(popx, popy, main = "Power vs Sample Size", xlab = "Sample Size", ylab = "Power")

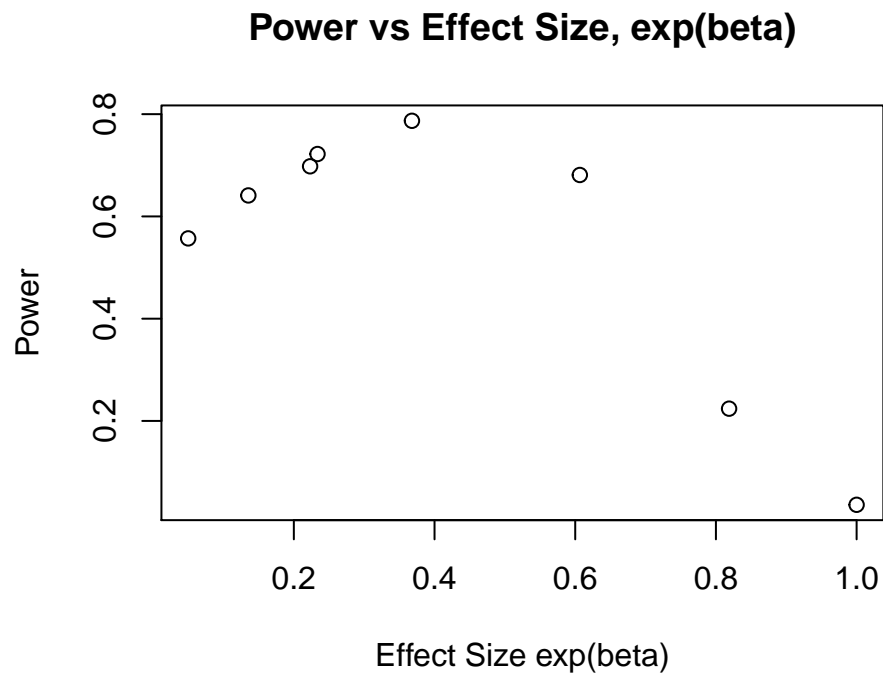
```



```
betax = c(0, -0.2, -0.5, -1, -1.454, -1.5, -2, -3)
betay = c(0.036, 0.224, 0.681, 0.787, 0.722, 0.698, 0.641, 0.557)
plot(betax, betay, main = "Power vs Effect Size (beta)", xlab = "Effect Size beta", ylab = "Power")
```



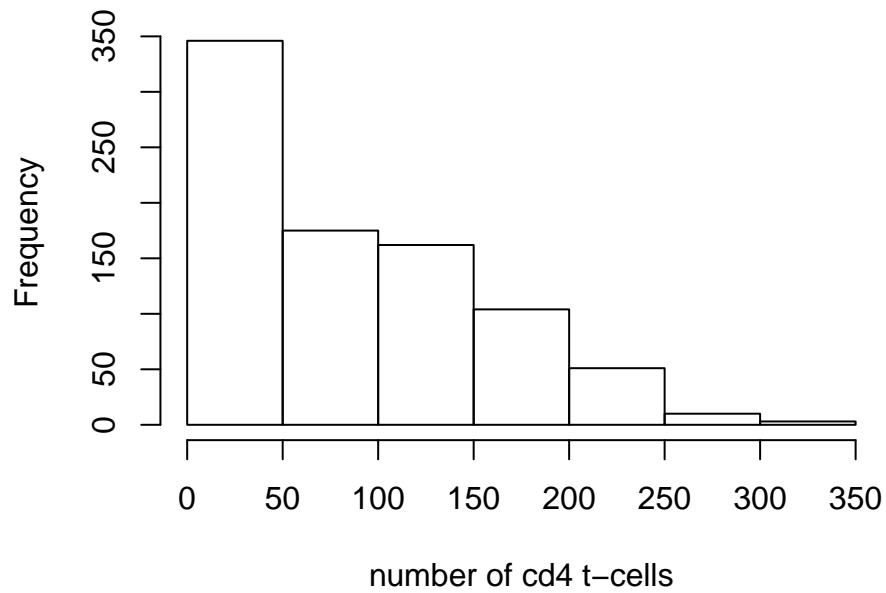
```
expbetax = c(1.00000000, 0.81873075, 0.60653066, 0.36787944, 0.23363388, 0.22313016, 0.13533528, 0.04978707)
plot(expbetax, betay, main = "Power vs Effect Size, exp(beta)", xlab = "Effect Size exp(beta)", ylab = "Power")
```



Some exploratory plots:

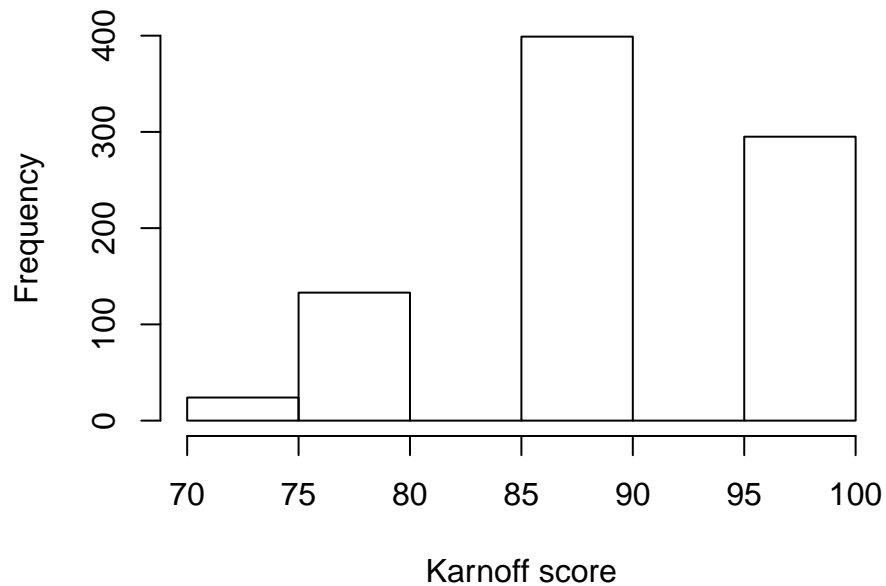
```
hist(AD$cd4, main = "Histogram of CD4 counts", xlab = "number of cd4 t-cells")
```

Histogram of CD4 counts



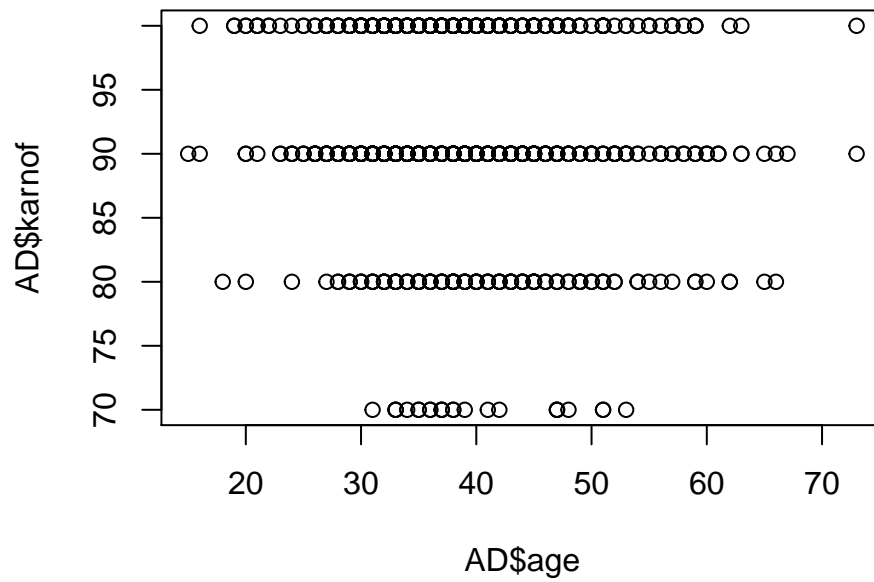
```
hist(AD$karnof, main = "Histogram of Karnof scores", xlab = "Karnoff score", breaks = 5)
```

Histogram of Karnof scores



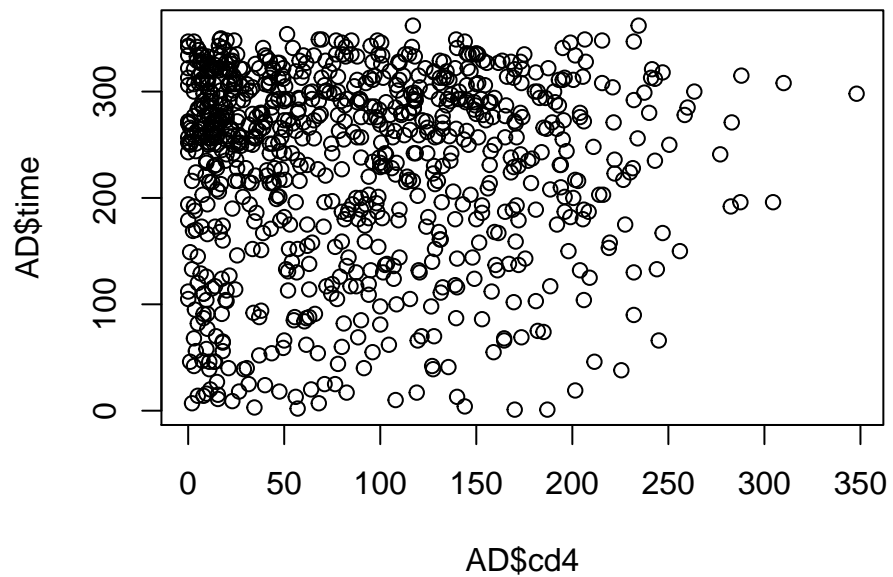
```
#I'm having a hard time getting the histogram breaks right - karnoff scores are multiples of 10 and it  
plot(AD$age, AD$karnof, main = "plot of karnof score vs age")
```

plot of karnof score vs age



```
plot(AD$cd4, AD$time, main = "time vs cd4 count, ignoring censoring")
```

time vs cd4 count, ignoring censoring



Working on COX model

Variables: Response: time, censor Not using: time.d, censor.d (because we're using "aids defining event or death", not just death) Explanatory: tx, txgrp, strat2, sex, raceth, ivdrug, hemophil, karnof, cd4, priorzdv, age

Let's check if we need interaction. Likelihood ratio test is $2\ln(\text{Likelihood of full}) - 2\ln(\text{Likelihood of reduced})$

```
full = coxph(Surv(time, censor) ~ (tx+strat2+sex+raceth+ivdrug+hemophil+karnof+cd4+priorzdv+age)^2, data = AD)
full$loglik[2]
```

```
## [1] -380.5977
```

```
red = coxph(Surv(time, censor) ~ tx+strat2+sex+raceth+ivdrug+hemophil+karnof+cd4+priorzdv+age, data = AD)
red$loglik[2]
```

```
## [1] -411.5144
```

```
length(red$coefficients)
```

```
## [1] 10
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2], df = length(full$coefficients) - length(red$coefficients))
```

```
## [1] 0.04845761
```

```
full
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ (tx + strat2 + sex + raceth +  
##      ivdrug + hemophil + karnof + cd4 + priorzdv + age)^2, data = AD)
```

```
##
```

	coef	exp(coef)	se(coef)	z	p
## tx	-5.501e+00	4.084e-03	4.369e+00	-1.259	0.20807
## strat2	-9.622e-01	3.820e-01	7.284e+00	-0.132	0.89490
## sex	1.137e+01	8.710e+04	5.589e+00	2.035	0.04182
## raceth	7.870e+00	2.617e+03	2.418e+00	3.255	0.00113
## ivdrug	-7.581e+00	5.101e-04	3.252e+00	-2.331	0.01976
## hemophil	1.996e+01	4.646e+08	2.107e+04	0.001	0.99924
## karnof	1.860e-01	1.204e+00	1.246e-01	1.493	0.13539
## cd4	-1.073e-01	8.983e-01	8.413e-02	-1.275	0.20221
## priorzdv	6.378e-02	1.066e+00	8.320e-02	0.767	0.44336
## age	4.436e-01	1.558e+00	2.030e-01	2.185	0.02890
## tx:strat2	1.039e-02	1.010e+00	9.648e-01	0.011	0.99141
## tx:sex	1.969e+00	7.163e+00	7.915e-01	2.487	0.01287
## tx:raceth	-6.084e-02	9.410e-01	3.554e-01	-0.171	0.86408
## tx:ivdrug	9.461e-02	1.099e+00	5.560e-01	0.170	0.86488
## tx:hemophil	-1.611e+01	1.008e-07	4.131e+03	-0.004	0.99689
## tx:karnof	2.166e-02	1.022e+00	3.952e-02	0.548	0.58369
## tx:cd4	5.997e-03	1.006e+00	1.114e-02	0.538	0.59049
## tx:priorzdv	-3.839e-02	9.623e-01	1.685e-02	-2.278	0.02273
## tx:age	3.409e-02	1.035e+00	3.266e-02	1.044	0.29656
## strat2:sex	-2.387e-01	7.876e-01	1.326e+00	-0.180	0.85716
## strat2:raceth	-4.696e-01	6.253e-01	6.182e-01	-0.760	0.44747
## strat2:ivdrug	5.624e-01	1.755e+00	7.919e-01	0.710	0.47756
## strat2:hemophil	1.885e+00	6.585e+00	3.504e+00	0.538	0.59061
## strat2:karnof	-1.333e-02	9.868e-01	5.889e-02	-0.226	0.82091
## strat2:cd4	-2.740e-03	9.973e-01	1.214e-02	-0.226	0.82142

```
## strat2:priorzdv      4.170e-02  1.043e+00  2.261e-02  1.845  0.06510
## strat2:age           4.497e-02  1.046e+00  4.961e-02  0.906  0.36468
## sex:raceth          -1.886e+00  1.516e-01  6.562e-01 -2.875  0.00404
## sex:ivdrug           4.405e-01  1.553e+00  6.784e-01  0.649  0.51617
## sex:hemophil        -2.067e+01  1.060e-09  2.085e+04 -0.001  0.99921
## sex:karnof          -7.926e-02  9.238e-01  5.430e-02 -1.460  0.14439
## sex:cd4             -8.461e-03  9.916e-01  1.769e-02 -0.478  0.63250
## sex:priorzdv        -3.032e-04  9.997e-01  2.286e-02 -0.013  0.98941
## sex:age            -5.357e-02  9.478e-01  4.356e-02 -1.230  0.21880
## raceth:ivdrug        2.133e-01  1.238e+00  2.426e-01  0.879  0.37933
## raceth:hemophil     -1.350e+01  1.373e-06  3.044e+03 -0.004  0.99646
## raceth:karnof       -5.435e-02  9.471e-01  2.404e-02 -2.261  0.02374
## raceth:cd4          7.225e-03  1.007e+00  6.483e-03  1.114  0.26509
## raceth:priorzdv     -3.059e-03  9.969e-01  9.559e-03 -0.320  0.74898
## raceth:age          -3.873e-02  9.620e-01  2.056e-02 -1.884  0.05957
## ivdrug:hemophil      2.633e+01  2.712e+11  6.088e+03  0.004  0.99655
## ivdrug:karnof        5.604e-02  1.058e+00  2.796e-02  2.004  0.04506
## ivdrug:cd4          -1.876e-02  9.814e-01  1.333e-02 -1.408  0.15921
## ivdrug:priorzdv      2.655e-02  1.027e+00  1.055e-02  2.517  0.01182
## ivdrug:age           2.986e-02  1.030e+00  3.208e-02  0.931  0.35186
## hemophil:karnof     -8.703e-02  9.167e-01  1.461e-01 -0.596  0.55134
## hemophil:cd4        -9.746e-03  9.903e-01  4.981e-02 -0.196  0.84486
## hemophil:priorzdv   -4.504e-02  9.560e-01  3.127e-02 -1.440  0.14975
## hemophil:age        -7.388e-02  9.288e-01  1.262e-01 -0.586  0.55820
## karnof:cd4           1.096e-03  1.001e+00  6.856e-04  1.599  0.10983
## karnof:priorzdv     -4.181e-04  9.996e-01  7.074e-04 -0.591  0.55454
## karnof:age          -4.056e-03  9.960e-01  2.148e-03 -1.888  0.05905
## cd4:priorzdv        -2.991e-04  9.997e-01  2.646e-04 -1.130  0.25830
## cd4:age              3.926e-04  1.000e+00  5.166e-04  0.760  0.44732
## priorzdv:age        -1.207e-03  9.988e-01  8.244e-04 -1.464  0.14323
##
## Likelihood ratio test=144.1 on 55 df, p=6.536e-10
## n= 851, number of events= 69
```

This is close enough to 0.05 that I'm not comfortable eliminating interaction.

lets eliminate hemophil entirely

```
full = coxph(Surv(time, censor) ~ (tx+strat2+sex+raceth+ivdrug+hemophil+karnof+cd4+priorzdv+age)^2, data = AD)
full$loglik[2]

## [1] -380.5977

red = coxph(Surv(time, censor) ~ (tx+strat2+sex+raceth+ivdrug+karnof+cd4+priorzdv+age)^2, data = AD)
red$loglik[2]

## [1] -384.3829

1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2] , df = length(full$coefficients) - length(red$coefficients))

## [1] 0.6707304

#red
```

We can eliminate hemophil entirely. Let's try that with strat 2

```
full = coxph(Surv(time, censor) ~ (tx+strat2+sex+raceth+ivdrug+karnof+cd4+priorzdv+age)^2, data = AD)
full$loglik[2]
```



```
## [1] -384.3829
```

```
red = coxph(Surv(time, censor) ~ (tx+sex+raceth+ivdrug+karnof+cd4+priorzdv+age)^2, data = AD)
red$loglik[2]
```

```
## [1] -387.8855
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2] , df = length(full$coefficients) - length(red$coefficients))
```

```
## [1] 0.6365669
```

```
red
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ (tx + sex + raceth + ivdrug +  
## karnof + cd4 + priorzdv + age)^2, data = AD)
```

```
##
```

	coef	exp(coef)	se(coef)	z	p
## tx	-6.310e+00	1.818e-03	4.296e+00	-1.469	0.14191
## sex	1.144e+01	9.324e+04	5.571e+00	2.054	0.03999
## raceth	7.255e+00	1.415e+03	2.301e+00	3.153	0.00162
## ivdrug	-6.516e+00	1.480e-03	2.981e+00	-2.186	0.02881
## karnof	1.679e-01	1.183e+00	1.184e-01	1.418	0.15611
## cd4	-7.608e-02	9.267e-01	5.281e-02	-1.441	0.14969
## priorzdv	-3.132e-03	9.969e-01	6.658e-02	-0.047	0.96248
## age	4.069e-01	1.502e+00	1.893e-01	2.150	0.03156
## tx:sex	2.125e+00	8.372e+00	7.668e-01	2.771	0.00558
## tx:raceth	1.504e-01	1.162e+00	3.420e-01	0.440	0.66003
## tx:ivdrug	4.191e-02	1.043e+00	5.375e-01	0.078	0.93784
## tx:karnof	2.505e-02	1.025e+00	3.819e-02	0.656	0.51179
## tx:cd4	4.788e-03	1.005e+00	7.088e-03	0.676	0.49930
## tx:priorzdv	-3.927e-02	9.615e-01	1.679e-02	-2.339	0.01933
## tx:age	3.413e-02	1.035e+00	3.090e-02	1.104	0.26946
## sex:raceth	-1.675e+00	1.874e-01	6.383e-01	-2.624	0.00870
## sex:ivdrug	3.202e-01	1.377e+00	6.562e-01	0.488	0.62555
## sex:karnof	-7.869e-02	9.243e-01	5.442e-02	-1.446	0.14816
## sex:cd4	-1.099e-02	9.891e-01	1.298e-02	-0.847	0.39703
## sex:priorzdv	-4.393e-03	9.956e-01	2.031e-02	-0.216	0.82876
## sex:age	-6.038e-02	9.414e-01	4.138e-02	-1.459	0.14451
## raceth:ivdrug	1.385e-03	1.001e+00	2.362e-01	0.006	0.99532
## raceth:karnof	-5.064e-02	9.506e-01	2.266e-02	-2.235	0.02543
## raceth:cd4	2.794e-03	1.003e+00	4.207e-03	0.664	0.50658
## raceth:priorzdv	-1.678e-03	9.983e-01	8.282e-03	-0.203	0.83939
## raceth:age	-3.409e-02	9.665e-01	1.899e-02	-1.795	0.07264
## ivdrug:karnof	5.035e-02	1.052e+00	2.597e-02	1.939	0.05256
## ivdrug:cd4	-9.323e-03	9.907e-01	7.417e-03	-1.257	0.20874
## ivdrug:priorzdv	2.362e-02	1.024e+00	9.567e-03	2.469	0.01356
## ivdrug:age	2.753e-02	1.028e+00	3.120e-02	0.883	0.37747
## karnof:cd4	6.475e-04	1.001e+00	4.480e-04	1.445	0.14839
## karnof:priorzdv	-2.237e-05	1.000e+00	6.636e-04	-0.034	0.97311
## karnof:age	-3.687e-03	9.963e-01	1.942e-03	-1.898	0.05769
## cd4:priorzdv	-4.921e-05	1.000e+00	1.366e-04	-0.360	0.71869
## cd4:age	4.985e-04	1.000e+00	3.177e-04	1.569	0.11670
## priorzdv:age	-2.041e-04	9.998e-01	5.395e-04	-0.378	0.70521

```
##
```

```
## Likelihood ratio test=129.5 on 36 df, p=1.777e-12
```

```
## n= 851, number of events= 69
```

Okay let's eliminate anything with a p-value of 0.1 or larger.

```
full = coxph(Surv(time, censor) ~ (tx+sex+raceth+ivdrug+karnof+cd4+priorzdv+age)^2, data = AD)
full$loglik[2]
```

```
## [1] -387.8855
```

```
red = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*(sex + karnof + age) + ivdrug*(karnof + p
red$loglik[2]
```

```
## [1] -395.5924
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2] , df = length(full$coefficients) - length(red$coeffic
```

```
## [1] 0.7522683
```

```
red
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ tx * (sex + priorzdv) +
##      raceth * (sex + karnof + age) + ivdrug * (karnof + priorzdv) +
##      karnof * age + cd4, data = AD)
```

```
##
```

	coef	exp(coef)	se(coef)	z	p
## tx	-1.773877	0.169674	0.886051	-2.002	0.04528
## sex	2.217867	9.187711	0.876329	2.531	0.01138
## priorzdv	-0.013380	0.986709	0.010620	-1.260	0.20770
## raceth	5.731699	308.492931	1.817565	3.154	0.00161
## karnof	0.027056	1.027426	0.081709	0.331	0.74055
## age	0.195550	1.215980	0.163624	1.195	0.23204
## ivdrug	-2.935066	0.053127	1.896507	-1.548	0.12171
## cd4	-0.016165	0.983965	0.003300	-4.899	9.63e-07
## tx:sex	1.644482	5.178329	0.681884	2.412	0.01588
## tx:priorzdv	-0.035672	0.964957	0.014254	-2.503	0.01233
## sex:raceth	-1.480031	0.227631	0.588021	-2.517	0.01184
## raceth:karnof	-0.036726	0.963940	0.019241	-1.909	0.05629
## raceth:age	-0.028096	0.972295	0.017905	-1.569	0.11660
## karnof:ivdrug	0.025423	1.025749	0.021618	1.176	0.23958
## priorzdv:ivdrug	0.016103	1.016233	0.006869	2.344	0.01906
## karnof:age	-0.001477	0.998524	0.001803	-0.819	0.41276

```
##
```

```
## Likelihood ratio test=114.1 on 16 df, p=< 2.2e-16
```

```
## n= 851, number of events= 69
```

We can do that, so let's eliminate karnof*age and karnof*ivdrug

```
full = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*(sex + karnof + age) + ivdrug*(karnof + p
full$loglik[2]
```

```
## [1] -395.5924
```

```
red = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*(sex + karnof + age) + ivdrug*priorzdv + c
red$loglik[2]
```

```
## [1] -396.5796
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2] , df = length(full$coefficients) - length(red$coeffic
```

```
## [1] 0.3726069
```

```
red
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ tx * (sex + priorzdv) +  
##      raceth * (sex + karnof + age) + ivdrug * priorzdv + cd4,  
##      data = AD)
```

```
##
```

	coef	exp(coef)	se(coef)	z	p
## tx	-1.651633	0.191736	0.877322	-1.883	0.05976
## sex	2.340187	10.383173	0.875181	2.674	0.00750
## priorzdv	-0.013415	0.986674	0.010817	-1.240	0.21490
## raceth	5.275991	195.584220	1.762487	2.993	0.00276
## karnof	-0.006901	0.993123	0.030802	-0.224	0.82274
## age	0.064565	1.066695	0.031193	2.070	0.03847
## ivdrug	-0.804226	0.447434	0.354498	-2.269	0.02329
## cd4	-0.015752	0.984371	0.003258	-4.834	1.34e-06
## tx:sex	1.551639	4.719197	0.674424	2.301	0.02141
## tx:priorzdv	-0.035633	0.964994	0.014349	-2.483	0.01302
## sex:raceth	-1.504677	0.222089	0.586714	-2.565	0.01033
## raceth:karnof	-0.031977	0.968528	0.019002	-1.683	0.09240
## raceth:age	-0.025912	0.974421	0.017915	-1.446	0.14806
## priorzdv:ivdrug	0.015861	1.015988	0.007212	2.199	0.02786

```
##
```

```
## Likelihood ratio test=112.1 on 14 df, p=< 2.2e-16
```

```
## n= 851, number of events= 69
```

Let's try raceth*age

```
full = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*sex + raceth*karnof + raceth*age + ivdrug  
full$loglik[2]
```

```
## [1] -396.5796
```

```
red = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*sex + raceth*karnof + ivdrug*priorzdv + c  
red$loglik[2]
```

```
## [1] -398.9897
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2] , df = length(full$coefficients) - length(red$coeffic
```

```
## [1] 0.08981025
```

```
red
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ tx * (sex + priorzdv) +  
##      raceth * sex + raceth * karnof + ivdrug * priorzdv + cd4,  
##      data = AD)
```

```
##
```

	coef	exp(coef)	se(coef)	z	p
## tx	-1.555087	0.211171	0.864569	-1.799	0.07207
## sex	2.283299	9.808991	0.877912	2.601	0.00930
## priorzdv	-0.013985	0.986112	0.010731	-1.303	0.19251
## raceth	4.287019	72.749312	1.625738	2.637	0.00837
## karnof	-0.009523	0.990523	0.030620	-0.311	0.75581
## ivdrug	-0.837568	0.432762	0.351801	-2.381	0.01728
## cd4	-0.015009	0.985103	0.003152	-4.762	1.91e-06

```
## tx:sex          1.473724  4.365463  0.668784  2.204  0.02755
## tx:priorzdvdv  -0.034477  0.966111  0.014294 -2.412  0.01587
## sex:raceth     -1.480506  0.227523  0.587672 -2.519  0.01176
## raceth:karnof  -0.032507  0.968016  0.018972 -1.713  0.08663
## priorzdvdv:ivdrug 0.016117  1.016247  0.007143  2.256  0.02405
##
## Likelihood ratio test=107.3 on 12 df, p=< 2.2e-16
## n= 851, number of events= 69
```

Let's try raceth*karnof

```
full = coxph(Surv(time, censor) ~ tx*(sex + priorzdvdv) + raceth*sex + raceth*karnof + ivdrug*priorzdvdv + cd4, data = AD)
full$loglik[2]
```

```
## [1] -398.9897
```

```
red = coxph(Surv(time, censor) ~ tx*(sex + priorzdvdv) + raceth*sex + karnof + ivdrug*priorzdvdv + cd4, data = AD)
red$loglik[2]
```

```
## [1] -400.5166
```

```
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2], df = length(full$coefficients) - length(red$coefficients))
```

```
## [1] 0.08054651
```

```
red
```

```
## Call:
```

```
## coxph(formula = Surv(time, censor) ~ tx * (sex + priorzdvdv) +
##       raceth * sex + karnof + ivdrug * priorzdvdv + cd4, data = AD)
```

```
##
##              coef exp(coef) se(coef)      z      p
## tx            -1.453900  0.233657  0.862695 -1.685 0.091931
## sex            2.469187 11.812844  0.866826  2.849 0.004392
## priorzdvdv    -0.015350  0.984767  0.010419 -1.473 0.140691
## raceth         1.670770  5.316262  0.625436  2.671 0.007554
## karnof        -0.056287  0.945268  0.014591 -3.857 0.000115
## ivdrug        -0.855328  0.425144  0.337158 -2.537 0.011185
## cd4           -0.015267  0.984849  0.003145 -4.854 1.21e-06
## tx:sex         1.392238  4.023845  0.666413  2.089 0.036694
## tx:priorzdvdv  -0.034642  0.965951  0.014195 -2.441 0.014666
## sex:raceth     -1.599834  0.201930  0.578075 -2.768 0.005648
## priorzdvdv:ivdrug 0.017395  1.017547  0.006697  2.597 0.009395
##
```

```
## Likelihood ratio test=104.2 on 11 df, p=< 2.2e-16
```

```
## n= 851, number of events= 69
```

Okay all the interaction terms are now below 0.05 and the only p-values above 0.05 are variables that are also in interaction terms. So, I have a model. Let's check against the full model with all interaction just to be sure.

```
full = coxph(Surv(time, censor) ~ (tx+strat2+sex+raceth+ivdrug+hemophil+karnof+cd4+priorzdvdv+age)^2, data = AD)
red = coxph(Surv(time, censor) ~ tx*(sex + priorzdvdv) + raceth*sex + ivdrug*priorzdvdv + karnof + cd4, data = AD)
1 - pchisq(2 * full$loglik[2] - 2 * red$loglik[2], df = length(full$coefficients) - length(red$coefficients))
```

```
## [1] 0.6505501
```

Alright that works then.

The model is:

```
model = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*sex + ivdrug*priorzdv + karnof + cd4, data = AD)
model
```

```
## Call:
## coxph(formula = Surv(time, censor) ~ tx * (sex + priorzdv) +
##       raceth * sex + ivdrug * priorzdv + karnof + cd4, data = AD)
##
##               coef exp(coef) se(coef)      z      p
## tx           -1.453900  0.233657  0.862695 -1.685 0.091931
## sex           2.469187 11.812844  0.866826  2.849 0.004392
## priorzdv     -0.015350  0.984767  0.010419 -1.473 0.140691
## raceth        1.670770  5.316262  0.625436  2.671 0.007554
## ivdrug       -0.855328  0.425144  0.337158 -2.537 0.011185
## karnof       -0.056287  0.945268  0.014591 -3.857 0.000115
## cd4          -0.015267  0.984849  0.003145 -4.854 1.21e-06
## tx:sex        1.392238  4.023845  0.666413  2.089 0.036694
## tx:priorzdv   -0.034642  0.965951  0.014195 -2.441 0.014666
## sex:raceth    -1.599834  0.201930  0.578075 -2.768 0.005648
## priorzdv:ivdrug 0.017395  1.017547  0.006697  2.597 0.009395
##
## Likelihood ratio test=104.2 on 11 df, p< 2.2e-16
## n= 851, number of events= 69
```

Let's test removing tx from the model to see what that does to the p-value so I can talk about it in my results.

```
model = coxph(Surv(time, censor) ~ tx*(sex + priorzdv) + raceth*sex + ivdrug*priorzdv + karnof + cd4, data = AD)
modelNoTx = coxph(Surv(time, censor) ~ raceth*sex + ivdrug*priorzdv + karnof + cd4, data = AD)
1 - pchisq(2 * model$loglik[2] - 2 * modelNoTx$loglik[2], df = length(model$coefficients) - length(modelNoTx$coefficients))

## [1] 0.000307816
```

P value of 0.000307816 that model is the same without tx (so model is not the same)