

Stepwise2

Zelin Wang

2023-09-09

Load data

```
cs_data<-read.csv("CS.csv")
```

Set random seed for subsequent random selection and assignment operations

```
set.seed(2023)
```

Partition data and create index matrix of selected values

```
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
index <- createDataPartition(cs_data$os, p=.8, list=FALSE, times=1)
```

Create test and train data frames

```
train_df <- cs_data[index,]  
test_df <- cs_data[-index,]
```

Verify number of rows (cases) in each data frame

```
nrow(train_df)
```

```
## [1] 15719
```

```
nrow(test_df)
```

```
## [1] 3929
```

```
#define response variable
```

```
y_nom <- train_df$os
```

```
#define matrix of predictor variables
```

```
x_nom <- data.matrix(train_df[, c(2:21)])  
data_nom <- data.frame(x_nom,y_nom)
```

Generate full model, exclude all cod related variables, and numeric variables.

```
full_model<-glm(y_nom~as.factor(gender)+as.factor(race)+as.factor(age)+as.factor(size)+as.factor(marry)+  
summary(full_model)
```

```
##
```

```
## Call:
```

```
## glm(formula = y_nom ~ as.factor(gender) + as.factor(race) + as.factor(age) +  
##   as.factor(size) + as.factor(marry) + as.factor(income) +  
##   as.factor(site) + as.factor(grade) + as.factor(kind) + as.factor(T) +  
##   as.factor(N) + as.factor(surgery_pri) + as.factor(RX_Summ) +  
##   as.factor(radiate) + as.factor(chem) + as.factor(CEA) + as.factor(bone) +  
##   as.factor(brain) + as.factor(lung) + as.factor(group), family = binomial(link = "logit"),  
##   data = data_nom)
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -2.1610  -0.5720  -0.3813  -0.2166   2.8853
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)  
## (Intercept)    -1.56599    0.29114  -5.379 7.50e-08 ***  
## as.factor(gender)2    -0.02130    0.04731  -0.450 0.652499  
## as.factor(race)2      0.04905    0.06469   0.758 0.448326  
## as.factor(race)3     -0.20023    0.08162  -2.453 0.014163 *  
## as.factor(age)2      -0.07145    0.17647  -0.405 0.685570  
## as.factor(age)3      0.38235    0.17389   2.199 0.027894 *  
## as.factor(size)2     0.19277    0.05835   3.304 0.000954 ***  
## as.factor(size)3     0.10832    0.06143   1.763 0.077829 .
```

```
## as.factor(marry)2      -0.17557    0.05915   -2.968 0.002998 **
## as.factor(marry)3       0.01998    0.06582    0.304 0.761425
## as.factor(income)2     -0.01414    0.06335   -0.223 0.823394
## as.factor(income)3     -0.04624    0.07231   -0.639 0.522546
## as.factor(site)2       0.30228    0.04868    6.210 5.31e-10 ***
## as.factor(grade)2      0.16418    0.13180    1.246 0.212914
## as.factor(grade)3      0.83737    0.13863    6.041 1.54e-09 ***
## as.factor(grade)4      0.75785    0.18684    4.056 4.99e-05 ***
## as.factor(grade)9      0.59052    0.13418    4.401 1.08e-05 ***
## as.factor(kind)2       0.32196    0.07975    4.037 5.41e-05 ***
## as.factor(T)2         -0.05885    0.08364   -0.704 0.481678
## as.factor(T)3          0.15454    0.07724    2.001 0.045409 *
## as.factor(N)2          0.08048    0.08985    0.896 0.370427
## as.factor(N)3          0.19448    0.06432    3.024 0.002497 **
## as.factor(surgery_pri)1 -0.42107    0.13861   -3.038 0.002382 **
## as.factor(RX_Summ)1    -0.07238    0.08051   -0.899 0.368655
## as.factor(radiate)1    -0.28655    0.08730   -3.282 0.001030 **
## as.factor(chem)1      -2.24326    0.04848  -46.272 < 2e-16 ***
## as.factor(CEA)1        0.47577    0.09554    4.980 6.36e-07 ***
## as.factor(CEA)2        0.39635    0.10076    3.934 8.36e-05 ***
## as.factor(bone)1       0.64492    0.08581    7.516 5.66e-14 ***
## as.factor(bone)2      -0.20963    0.22611   -0.927 0.353874
## as.factor(brain)1      1.21830    0.18503    6.584 4.57e-11 ***
## as.factor(brain)2     -0.24428    0.20752   -1.177 0.239126
## as.factor(lung)1       0.31745    0.05138    6.178 6.47e-10 ***
## as.factor(lung)2       0.09871    0.15915    0.620 0.535119
## as.factor(group)2      0.33871    0.10019    3.381 0.000723 ***
## as.factor(group)9      0.52823    0.14837    3.560 0.000371 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 16414 on 15718 degrees of freedom
## Residual deviance: 12389 on 15683 degrees of freedom
## AIC: 12461
##
## Number of Fisher Scoring iterations: 5
```

Perform stepwise regression

create null model

```
null_model<-glm(y_nom~1,data=data_nom,family=binomial(link="logit"))
```

Perform forward selection

set seed

```
set.seed(2023)
```

AIC threshold (Akaike Information Criterion)

```
model_step1 = step(null_model,  
                    scope = list(lower=null_model, upper=full_model),  
                    direction = "forward", k= 2, trace=0)
```

```
###summary model  
summary(model_step1)
```

```
##  
## Call:  
## glm(formula = y_nom ~ as.factor(chem) + as.factor(group) + as.factor(grade) +  
##      as.factor(age) + as.factor(bone) + as.factor(site) + as.factor(lung) +  
##      as.factor(brain) + as.factor(CEA) + as.factor(T) + as.factor(marry) +  
##      as.factor(kind) + as.factor(radiate) + as.factor(surgery_pri) +  
##      as.factor(size) + as.factor(N) + as.factor(race), family = binomial(link = "logit"),  
##      data = data_nom)  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max   
## -2.1650  -0.5723  -0.3821  -0.2159   2.8734   
##  
## Coefficients:  
##              Estimate Std. Error z value Pr(>|z|)      
## (Intercept)    -1.60286    0.28491  -5.626 1.85e-08 ***  
## as.factor(chem)1 -2.24607    0.04841 -46.398 < 2e-16 ***  
## as.factor(group)2  0.34018    0.10014  3.397 0.000681 ***  
## as.factor(group)9  0.53303    0.14815  3.598 0.000321 ***  
## as.factor(grade)2  0.16286    0.13178  1.236 0.216499   
## as.factor(grade)3  0.83634    0.13856  6.036 1.58e-09 ***  
## as.factor(grade)4  0.75880    0.18679  4.062 4.86e-05 ***  
## as.factor(grade)9  0.58808    0.13412  4.385 1.16e-05 ***  
## as.factor(age)2   -0.07351    0.17644  -0.417 0.676962   
## as.factor(age)3    0.38129    0.17382  2.194 0.028262 *   
## as.factor(bone)1   0.64492    0.08579  7.517 5.60e-14 ***  
## as.factor(bone)2  -0.21274    0.22617  -0.941 0.346903   
## as.factor(site)2   0.30507    0.04843  6.299 2.99e-10 ***  
## as.factor(lung)1   0.31849    0.05136  6.201 5.62e-10 ***  
## as.factor(lung)2   0.10072    0.15907  0.633 0.526605   
## as.factor(brain)1  1.20828    0.18437  6.553 5.62e-11 ***  
## as.factor(brain)2 -0.23986    0.20746  -1.156 0.247600   
## as.factor(CEA)1    0.47745    0.09554  4.998 5.80e-07 ***  
## as.factor(CEA)2    0.40026    0.10071  3.974 7.06e-05 ***  
## as.factor(T)2     -0.06241    0.08354  -0.747 0.455043
```

```
## as.factor(T)3          0.15145    0.07712    1.964 0.049572 *
## as.factor(marry)2      -0.17587    0.05911   -2.975 0.002927 **
## as.factor(marry)3      0.02573    0.06495    0.396 0.691955
## as.factor(kind)2       0.32114    0.07970    4.029 5.60e-05 ***
## as.factor(radiate)1    -0.28760    0.08729   -3.295 0.000984 ***
## as.factor(surgery_pri)1 -0.42867    0.13813   -3.103 0.001913 **
## as.factor(size)2       0.19163    0.05832    3.286 0.001016 **
## as.factor(size)3       0.10767    0.06139    1.754 0.079436 .
## as.factor(N)2          0.08008    0.08982    0.892 0.372611
## as.factor(N)3          0.19288    0.06428    3.001 0.002695 **
## as.factor(race)2       0.05383    0.06415    0.839 0.401373
## as.factor(race)3      -0.20776    0.08052   -2.580 0.009876 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 16414  on 15718  degrees of freedom
## Residual deviance: 12390  on 15687  degrees of freedom
## AIC: 12454
##
## Number of Fisher Scoring iterations: 5
```

```
cat("AIC = ", AIC(model_step1), sep = "")
```

```
## AIC = 12454.4
```

BIC threshold (Bayesian information criterion)

```
model_step2 = step(null_model,
  scope = list(lower=null_model, upper=full_model),
  direction = "forward", k= log(nrow(cs_data)), trace=0)
```

```
#summary model
summary(model_step2)
```

```
##
## Call:
## glm(formula = y_nom ~ as.factor(chem) + as.factor(group) + as.factor(grade) +
##      as.factor(age) + as.factor(bone) + as.factor(site) + as.factor(lung) +
##      as.factor(brain) + as.factor(CEA) + as.factor(kind) + as.factor(surgery_pri) +
##      as.factor(radiate) + as.factor(marry), family = binomial(link = "logit"),
##      data = data_nom)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1223  -0.5675  -0.3870  -0.2172   2.8491
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.51601    0.27356  -5.542 2.99e-08 ***
```

```
## as.factor(chem)1      -2.24321    0.04817 -46.568 < 2e-16 ***
## as.factor(group)2     0.39888    0.07879   5.063 4.13e-07 ***
## as.factor(group)9     0.63650    0.14449   4.405 1.06e-05 ***
## as.factor(grade)2     0.15781    0.13162   1.199 0.230563
## as.factor(grade)3     0.83912    0.13829   6.068 1.30e-09 ***
## as.factor(grade)4     0.76295    0.18638   4.094 4.25e-05 ***
## as.factor(grade)9     0.63033    0.13368   4.715 2.41e-06 ***
## as.factor(age)2       -0.06491    0.17650  -0.368 0.713041
## as.factor(age)3       0.39107    0.17392   2.249 0.024540 *
## as.factor(bone)1      0.65077    0.08547   7.614 2.66e-14 ***
## as.factor(bone)2     -0.18704    0.22560  -0.829 0.407061
## as.factor(site)2      0.32403    0.04804   6.746 1.52e-11 ***
## as.factor(lung)1      0.32607    0.05118   6.371 1.87e-10 ***
## as.factor(lung)2      0.11808    0.15872   0.744 0.456907
## as.factor(brain)1     1.20730    0.18418   6.555 5.57e-11 ***
## as.factor(brain)2    -0.19486    0.20655  -0.943 0.345467
## as.factor(CEA)1       0.47545    0.09530   4.989 6.08e-07 ***
## as.factor(CEA)2       0.41365    0.10036   4.122 3.76e-05 ***
## as.factor(kind)2      0.33772    0.07938   4.254 2.10e-05 ***
## as.factor(surgery_pri)1 -0.49516    0.13676  -3.621 0.000294 ***
## as.factor(radiate)1   -0.31225    0.08670  -3.602 0.000316 ***
## as.factor(marry)2     -0.19756    0.05813  -3.399 0.000677 ***
## as.factor(marry)3      0.01491    0.06453   0.231 0.817285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 16414  on 15718  degrees of freedom
## Residual deviance: 12434  on 15695  degrees of freedom
## AIC: 12482
##
## Number of Fisher Scoring iterations: 5
```

```
cat("BIC = ", BIC(model_step2), sep = "")
```

```
## BIC = 12665.69
```

Perform backward elimination

AIC threshold (Akaike Information Criterion)

```
model_step3 = step(full_model,
                    direction = "backward", k= 2, trace=0)
```

```
#summary model
summary(model_step3)
```

```
##
## Call:
## glm(formula = y_nom ~ as.factor(race) + as.factor(age) + as.factor(size) +
```

```

##      as.factor(marry) + as.factor(site) + as.factor(grade) + as.factor(kind) +
##      as.factor(T) + as.factor(N) + as.factor(surgery_pri) + as.factor(radiate) +
##      as.factor(chem) + as.factor(CEA) + as.factor(bone) + as.factor(brain) +
##      as.factor(lung) + as.factor(group), family = binomial(link = "logit"),
##      data = data_nom)
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -2.1650   -0.5723   -0.3821   -0.2159    2.8734
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.60286    0.28491  -5.626 1.85e-08 ***
## as.factor(race)2     0.05383    0.06415   0.839 0.401373
## as.factor(race)3    -0.20776    0.08052  -2.580 0.009876 **
## as.factor(age)2     -0.07351    0.17644  -0.417 0.676962
## as.factor(age)3     0.38129    0.17382   2.194 0.028262 *
## as.factor(size)2     0.19163    0.05832   3.286 0.001016 **
## as.factor(size)3     0.10767    0.06139   1.754 0.079436 .
## as.factor(marry)2   -0.17587    0.05911  -2.975 0.002927 **
## as.factor(marry)3    0.02573    0.06495   0.396 0.691955
## as.factor(site)2    0.30507    0.04843   6.299 2.99e-10 ***
## as.factor(grade)2    0.16286    0.13178   1.236 0.216499
## as.factor(grade)3    0.83634    0.13856   6.036 1.58e-09 ***
## as.factor(grade)4    0.75880    0.18679   4.062 4.86e-05 ***
## as.factor(grade)9    0.58808    0.13412   4.385 1.16e-05 ***
## as.factor(kind)2     0.32114    0.07970   4.029 5.60e-05 ***
## as.factor(T)2       -0.06241    0.08354  -0.747 0.455043
## as.factor(T)3        0.15145    0.07712   1.964 0.049572 *
## as.factor(N)2        0.08008    0.08982   0.892 0.372611
## as.factor(N)3        0.19288    0.06428   3.001 0.002695 **
## as.factor(surgery_pri)1 -0.42867    0.13813  -3.103 0.001913 **
## as.factor(radiate)1  -0.28760    0.08729  -3.295 0.000984 ***
## as.factor(chem)1     -2.24607    0.04841 -46.398 < 2e-16 ***
## as.factor(CEA)1      0.47745    0.09554   4.998 5.80e-07 ***
## as.factor(CEA)2      0.40026    0.10071   3.974 7.06e-05 ***
## as.factor(bone)1     0.64492    0.08579   7.517 5.60e-14 ***
## as.factor(bone)2    -0.21274    0.22617  -0.941 0.346903
## as.factor(brain)1    1.20828    0.18437   6.553 5.62e-11 ***
## as.factor(brain)2   -0.23986    0.20746  -1.156 0.247600
## as.factor(lung)1     0.31849    0.05136   6.201 5.62e-10 ***
## as.factor(lung)2     0.10072    0.15907   0.633 0.526605
## as.factor(group)2    0.34018    0.10014   3.397 0.000681 ***
## as.factor(group)9    0.53303    0.14815   3.598 0.000321 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 16414  on 15718  degrees of freedom
## Residual deviance: 12390  on 15687  degrees of freedom
## AIC: 12454
##
## Number of Fisher Scoring iterations: 5

```

```
cat("AIC = ", AIC(model_step3), sep = "")
```

```
## AIC = 12454.4
```

BIC threshold (Bayesian information criterion)

```
model_step4 = step(full_model,  
                    direction = "backward", k= log(nrow(cs_data)), trace=0)
```

```
#summary model  
summary(model_step4)
```

```
##  
## Call:  
## glm(formula = y_nom ~ as.factor(age) + as.factor(marry) + as.factor(site) +  
##   as.factor(grade) + as.factor(kind) + as.factor(surgery_pri) +  
##   as.factor(radiate) + as.factor(chem) + as.factor(CEA) + as.factor(bone) +  
##   as.factor(brain) + as.factor(lung) + as.factor(group), family = binomial(link = "logit"),  
##   data = data_nom)  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max   
## -2.1223  -0.5675  -0.3870  -0.2172   2.8491   
##  
## Coefficients:  
##              Estimate Std. Error z value Pr(>|z|)      
## (Intercept)    -1.51601    0.27356  -5.542 2.99e-08 ***  
## as.factor(age)2    -0.06491    0.17650  -0.368 0.713041   
## as.factor(age)3     0.39107    0.17392   2.249 0.024540 *   
## as.factor(marry)2  -0.19756    0.05813  -3.399 0.000677 ***  
## as.factor(marry)3   0.01491    0.06453   0.231 0.817285   
## as.factor(site)2    0.32403    0.04804   6.746 1.52e-11 ***  
## as.factor(grade)2   0.15781    0.13162   1.199 0.230563   
## as.factor(grade)3   0.83912    0.13829   6.068 1.30e-09 ***  
## as.factor(grade)4   0.76295    0.18638   4.094 4.25e-05 ***  
## as.factor(grade)9   0.63033    0.13368   4.715 2.41e-06 ***  
## as.factor(kind)2    0.33772    0.07938   4.254 2.10e-05 ***  
## as.factor(surgery_pri)1 -0.49516    0.13676  -3.621 0.000294 ***  
## as.factor(radiate)1  -0.31225    0.08670  -3.602 0.000316 ***  
## as.factor(chem)1    -2.24321    0.04817 -46.568 < 2e-16 ***  
## as.factor(CEA)1     0.47545    0.09530   4.989 6.08e-07 ***  
## as.factor(CEA)2     0.41365    0.10036   4.122 3.76e-05 ***  
## as.factor(bone)1     0.65077    0.08547   7.614 2.66e-14 ***  
## as.factor(bone)2    -0.18704    0.22560  -0.829 0.407061   
## as.factor(brain)1    1.20730    0.18418   6.555 5.57e-11 ***  
## as.factor(brain)2   -0.19486    0.20655  -0.943 0.345467   
## as.factor(lung)1     0.32607    0.05118   6.371 1.87e-10 ***  
## as.factor(lung)2     0.11808    0.15872   0.744 0.456907   
## as.factor(group)2    0.39888    0.07879   5.063 4.13e-07 ***  
## as.factor(group)9    0.63650    0.14449   4.405 1.06e-05 ***
```



```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 16414  on 15718  degrees of freedom
## Residual deviance: 12434  on 15695  degrees of freedom
## AIC: 12482
##
## Number of Fisher Scoring iterations: 5
```

```
cat("bIC = ", BIC(model_step3), sep = "")
```

```
## bIC = 12699.61
```

Perform stepwise selection (both)

AIC threshold (Akaike Information Criterion)

```
model_step5 = step(null_model,
                    scope = list(lower=null_model, upper=full_model),
                    direction = "both", k= 2, trace=0)
```

```
#summary model
summary(model_step5)
```

```
##
## Call:
## glm(formula = y_nom ~ as.factor(chem) + as.factor(group) + as.factor(grade) +
##      as.factor(age) + as.factor(bone) + as.factor(site) + as.factor(lung) +
##      as.factor(brain) + as.factor(CEA) + as.factor(T) + as.factor(marry) +
##      as.factor(kind) + as.factor(radiate) + as.factor(surgery_pri) +
##      as.factor(size) + as.factor(N) + as.factor(race), family = binomial(link = "logit"),
##      data = data_nom)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1650  -0.5723  -0.3821  -0.2159   2.8734
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.60286    0.28491  -5.626 1.85e-08 ***
## as.factor(chem)1 -2.24607    0.04841 -46.398 < 2e-16 ***
## as.factor(group)2  0.34018    0.10014   3.397 0.000681 ***
## as.factor(group)9  0.53303    0.14815   3.598 0.000321 ***
## as.factor(grade)2  0.16286    0.13178   1.236 0.216499
## as.factor(grade)3  0.83634    0.13856   6.036 1.58e-09 ***
## as.factor(grade)4  0.75880    0.18679   4.062 4.86e-05 ***
## as.factor(grade)9  0.58808    0.13412   4.385 1.16e-05 ***
## as.factor(age)2   -0.07351    0.17644  -0.417 0.676962
## as.factor(age)3    0.38129    0.17382   2.194 0.028262 *
```

```
## as.factor(bone)1      0.64492    0.08579    7.517 5.60e-14 ***
## as.factor(bone)2     -0.21274    0.22617   -0.941 0.346903
## as.factor(site)2      0.30507    0.04843    6.299 2.99e-10 ***
## as.factor(lung)1      0.31849    0.05136    6.201 5.62e-10 ***
## as.factor(lung)2      0.10072    0.15907    0.633 0.526605
## as.factor(brain)1     1.20828    0.18437    6.553 5.62e-11 ***
## as.factor(brain)2    -0.23986    0.20746   -1.156 0.247600
## as.factor(CEA)1       0.47745    0.09554    4.998 5.80e-07 ***
## as.factor(CEA)2       0.40026    0.10071    3.974 7.06e-05 ***
## as.factor(T)2        -0.06241    0.08354   -0.747 0.455043
## as.factor(T)3         0.15145    0.07712    1.964 0.049572 *
## as.factor(marry)2     -0.17587    0.05911   -2.975 0.002927 **
## as.factor(marry)3      0.02573    0.06495    0.396 0.691955
## as.factor(kind)2      0.32114    0.07970    4.029 5.60e-05 ***
## as.factor(radiate)1   -0.28760    0.08729   -3.295 0.000984 ***
## as.factor(surgery_pri)1 -0.42867    0.13813   -3.103 0.001913 **
## as.factor(size)2      0.19163    0.05832    3.286 0.001016 **
## as.factor(size)3      0.10767    0.06139    1.754 0.079436 .
## as.factor(N)2         0.08008    0.08982    0.892 0.372611
## as.factor(N)3         0.19288    0.06428    3.001 0.002695 **
## as.factor(race)2      0.05383    0.06415    0.839 0.401373
## as.factor(race)3     -0.20776    0.08052   -2.580 0.009876 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 16414 on 15718 degrees of freedom
## Residual deviance: 12390 on 15687 degrees of freedom
## AIC: 12454
##
## Number of Fisher Scoring iterations: 5
```

```
cat("AIC = ", AIC(model_step5), sep = "")
```

```
## AIC = 12454.4
```

BIC threshold (Bayesian information criterion)

```
model_step6 = step(null_model,
  scope = list(lower=null_model, upper=full_model),
  direction = "both", k= log(nrow(cs_data)), trace=0)
```

```
#summary model
summary(model_step6)
```

```
##
## Call:
## glm(formula = y_nom ~ as.factor(chem) + as.factor(group) + as.factor(grade) +
##   as.factor(age) + as.factor(bone) + as.factor(site) + as.factor(lung) +
##   as.factor(brain) + as.factor(CEA) + as.factor(kind) + as.factor(surgery_pri) +
```

```

##      as.factor(radiate) + as.factor(marry), family = binomial(link = "logit"),
##      data = data_nom)
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -2.1223   -0.5675   -0.3870   -0.2172    2.8491
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.51601    0.27356  -5.542 2.99e-08 ***
## as.factor(chem)1    -2.24321    0.04817 -46.568 < 2e-16 ***
## as.factor(group)2     0.39888    0.07879   5.063 4.13e-07 ***
## as.factor(group)9     0.63650    0.14449   4.405 1.06e-05 ***
## as.factor(grade)2     0.15781    0.13162   1.199 0.230563
## as.factor(grade)3     0.83912    0.13829   6.068 1.30e-09 ***
## as.factor(grade)4     0.76295    0.18638   4.094 4.25e-05 ***
## as.factor(grade)9     0.63033    0.13368   4.715 2.41e-06 ***
## as.factor(age)2     -0.06491    0.17650  -0.368 0.713041
## as.factor(age)3      0.39107    0.17392   2.249 0.024540 *
## as.factor(bone)1      0.65077    0.08547   7.614 2.66e-14 ***
## as.factor(bone)2     -0.18704    0.22560  -0.829 0.407061
## as.factor(site)2      0.32403    0.04804   6.746 1.52e-11 ***
## as.factor(lung)1      0.32607    0.05118   6.371 1.87e-10 ***
## as.factor(lung)2      0.11808    0.15872   0.744 0.456907
## as.factor(brain)1     1.20730    0.18418   6.555 5.57e-11 ***
## as.factor(brain)2     -0.19486    0.20655  -0.943 0.345467
## as.factor(CEA)1       0.47545    0.09530   4.989 6.08e-07 ***
## as.factor(CEA)2       0.41365    0.10036   4.122 3.76e-05 ***
## as.factor(kind)2      0.33772    0.07938   4.254 2.10e-05 ***
## as.factor(surgery_pri)1 -0.49516    0.13676  -3.621 0.000294 ***
## as.factor(radiate)1    -0.31225    0.08670  -3.602 0.000316 ***
## as.factor(marry)2     -0.19756    0.05813  -3.399 0.000677 ***
## as.factor(marry)3      0.01491    0.06453   0.231 0.817285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 16414  on 15718  degrees of freedom
## Residual deviance: 12434  on 15695  degrees of freedom
## AIC: 12482
##
## Number of Fisher Scoring iterations: 5

cat("BIC = ", BIC(model_step6), sep = "")

## BIC = 12665.69

```