Homework 08
Joseph Blubaugh
jblubau1@tamu.edu
STAT 659-700

5.13

AIC shows that the reduced model with verw removed is the better model because the AIC is lower. Verw is highly insiginificant and so removing it results in better overall performance than leaving it in. However the likelihood ratio test shows that the reduced model is not significantly better than the full model.

5.9ad

a. This is a better test than the Chisquared Goodness of Fit test because for GOF you need each of the expected counts to be greater than 5 which will not happen in this case based on N and the number of levels.

b. `1 - pchisq(6.6, 6) = .359` The Hosmer-Lemeshow statistic follows a Chisquared distribution and in this case the model with the single term is insignificant and .359 would be the pvalue of the single term.

5.10c

A high pvalue indicates that weight is a good predictor of a female grab having satelite. If the pvalue were low (< .05) we would reject the null hypothesis that the model is a good fit. Since the pvalue is near one we conclude that the model is excellent fit.

```
library(ResourceSelection)

crabs = read.csv("crabs.csv")

mdl = glm(y ~ weight, family = binomial(), data = crabs)

hoslem.test(x = crabs$y, y = fitted(mdl), g = length(unique(crabs$group)))
```

```
    Hosmer and Lemeshow goodness of fit (GOF) test

data:  crabs$y, fitted(mdl)
X-squared = 1.5777, df = 6, p-value = 0.9542
```

5.14

a)

```
## Grouped
(grouped = data.frame(x = c(0, 1, 2), n = rep(4, 3), s = c(1, 2, 4)))
```

```
  x n s
1 0 4 1
2 1 4 2
3 2 4 4
```

```
## Ungrouped
(ungrouped = data.frame(x = c(rep(0, 4), rep(1, 4), rep(2, 4)),
  s = c(0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1)
))
```

```
   x s
1  0 0
2  0 0
3  0 0
4  0 1
5  1 0
6  1 0
7  1 1
8  1 1
9  2 1
10 2 1
11 2 1
12 2 1
```

```
## Grouped Data Intercept Only
(mdl.grouped.int = glm(cbind(s, n-s) ~ 1, family = binomial(), data = grouped))
```

```
Call:  glm(formula = cbind(s, n - s) ~ 1, family = binomial(), data = grouped)

Coefficients:
(Intercept)
     0.3365

Degrees of Freedom: 2 Total (i.e. Null);  2 Residual
Null Deviance:      6.257
Residual Deviance: 6.257    AIC: 11.94
```

```
## Grouped Data with X
(mdl.grouped.x = glm(cbind(s, n-s) ~ x, family = binomial(), data = grouped))
```

```
Call:  glm(formula = cbind(s, n - s) ~ x, family = binomial(), data = grouped)

Coefficients:
(Intercept)            x
     -1.503        2.060

Degrees of Freedom: 2 Total (i.e. Null);  1 Residual
Null Deviance:      6.257
Residual Deviance: 0.9844   AIC: 8.672
```

```
## Individual Observations Intercept Only
(mdl.ungrouped.int = glm(s ~ 1, family = binomial(), data = ungrouped))
```

```

Call:  glm(formula = s ~ 1, family = binomial(), data = ungrouped)

Coefficients:
(Intercept)
     0.3365

Degrees of Freedom: 11 Total (i.e. Null);  11 Residual
Null Deviance:      16.3
Residual Deviance: 16.3     AIC: 18.3
```

```
## Individual Observations with X
(mdl.ungrouped.x = glm(s ~ x, family = binomial(), data = ungrouped))
```

```

Call:  glm(formula = s ~ x, family = binomial(), data = ungrouped)

Coefficients:
(Intercept)            x
     -1.503        2.060

Degrees of Freedom: 11 Total (i.e. Null);  10 Residual
Null Deviance:      16.3
Residual Deviance: 11.03    AIC: 15.03
```

```
## Log Liklihood
logLik(mdl.grouped.int); logLik(mdl.grouped.x)
```

```
'log Lik.' -4.972265 (df=1)
```

```
'log Lik.' -2.336075 (df=2)
```

```
logLik(mdl.ungrouped.int); logLik(mdl.ungrouped.x)
```

```
'log Lik.' -8.150319 (df=1)
```

```
'log Lik.' -5.514129 (df=2)
```

b) The deviances are different because its based on the log likelihood which is different for all 4 models.

```
## Deviance for grouped data
anova(mdl.grouped.int, mdl.grouped.x)
```

```
Analysis of Deviance Table

Model 1: cbind(s, n - s) ~ 1
Model 2: cbind(s, n - s) ~ x
  Resid. Df Resid. Dev Df Deviance
1         2     6.2568
2         1     0.9844  1   5.2724
```

```
## Deviance for ungrouped data
anova(mdl.ungrouped.int, mdl.ungrouped.x)
```

```
Analysis of Deviance Table

Model 1: s ~ 1
Model 2: s ~ x
  Resid. Df Resid. Dev Df Deviance
1        11     16.301
2        10     11.028  1   5.2724
```

c) The difference in deviance is the same when comparing the 2 models in the different datasets even though the Log Likelihood is different for the 2 sets of models

5.16

Pvalue: `(1.8006e+02 - 1.9355e-04) = 180.06 > 5.99 = qchisq(.95, 2)`. We conclude that the model is highly significant due to the difference between the null and residual deviance which is much larger than the critical value on 2 degrees of freedom.

Table 1: data.frame: dta

Grad	No.Grad	Race	Gender
498	298	White	Female
878	747	White	Male
54	89	Black	Female
197	463	Black	Male

```
mdl = glm(cbind(dta$Grad, dta$No.Grad) ~ Race + Gender, family = binomial(), data = dta)
summary(mdl)
```

```
Call:
glm(formula = cbind(dta$Grad, dta$No.Grad) ~ Race + Gender, family = binomial(),
    data = dta)

Deviance Residuals:
       1          2          3          4
-0.004812   0.003270   0.011335  -0.005588

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.50161    0.10004  -5.014 5.33e-07 ***
RaceWhite    1.01547    0.08723  11.641  < 2e-16 ***
GenderMale  -0.35244    0.08044  -4.381 1.18e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1.8006e+02  on 3  degrees of freedom
Residual deviance: 1.9355e-04  on 1  degrees of freedom
AIC: 33.029

Number of Fisher Scoring iterations: 2
```

5.18

Table 2: Smoking on Lung Cancer by City

city	smoke	yes	no
c1	yes	126	100
c1	no	35	61
c2	yes	908	688
c2	no	497	807
c3	yes	913	747
c3	no	336	598
c4	yes	235	172
c4	no	58	121
c5	yes	402	308
c5	no	121	215
c6	yes	182	156
c6	no	72	98
c7	yes	60	99
c7	no	11	43
c8	yes	104	89
c8	no	21	36

a. Smoking Effect: `exp(.777) = 2.17`. The odds of a smoker developing cancer are 2.17 times greater than a non-smoker.

```
mdl = glm(cbind(dta$yes, dta$no) ~ city + smoke, family = binomial(), data = dta)
summary(mdl)
```

```

Call:
glm(formula = cbind(dta$yes, dta$no) ~ city + smoke, family = binomial(),
    data = dta)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.21781  -0.14842  -0.00012   0.16817   1.35470

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.548682   0.118022  -4.649 3.34e-06 ***
cityc2       0.055618   0.119570   0.465    0.642
cityc3      -0.027739   0.120071  -0.231    0.817
cityc4       0.005764   0.140911   0.041    0.967
cityc5       0.018187   0.129473   0.140    0.888
cityc6       0.028782   0.144755   0.199    0.842
cityc7      -0.745683   0.185519  -4.019 5.83e-05 ***
```

```
cityc8       -0.054906   0.170996  -0.321    0.748
smokeyes      0.777062   0.046775  16.613   < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 310.8951  on 15  degrees of freedom
Residual deviance:   5.1958  on  7  degrees of freedom
AIC: 121.05

Number of Fisher Scoring iterations: 3
```

b. The Pearson Chisquared statistic is smaller than the Critical value so we would conclude that the model is a good fit of the data

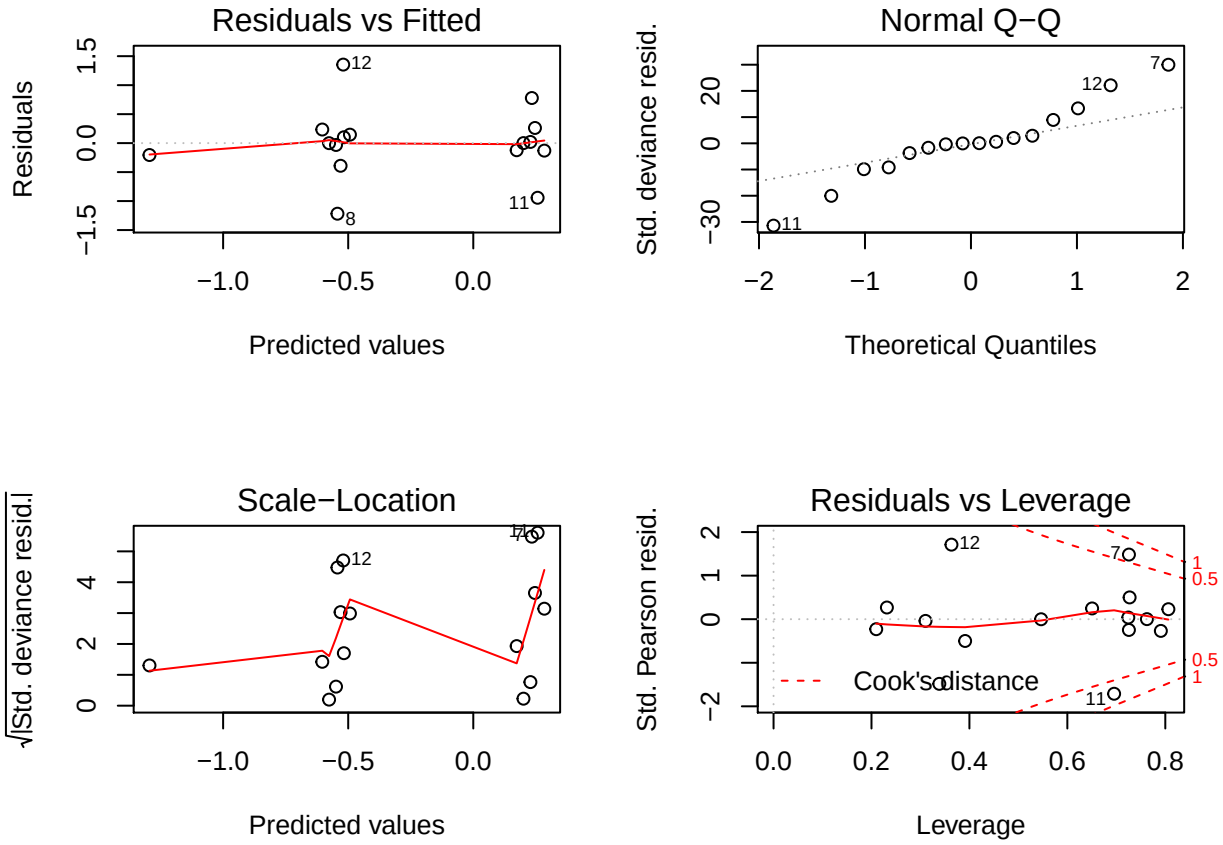
```
(Pearson.Chisq = sum(residuals(mdl, type = "pearson")^2))
```

```
[1] 5.199866
```

```
(Critical.Chisq = qchisq(.95, 7))
```

```
[1] 14.06714
```

c. The residual vs fitted and standardized residual plots both show no pattern in the residuals but the standardized residuals shows large deviance (> 3). The Normal QQ plot also shows large standardized residuals even though the plot is a relatively flat line. There are also a few observations which have high leverage and influence.. Removing points 7 and 11 may improve the model fit. We would conclude based on the diagnostic plots that the model is a not a good fit for the data.



5.19

Table 3: Admissions to Berkley

department	gender	yes	no
1	Male	512	313
1	Female	89	19
2	Male	353	207
2	Female	17	8
3	Male	120	205
3	Female	202	391
4	Male	138	279
4	Female	131	244
5	Male	53	138
5	Female	94	299
6	Male	22	351
6	Female	24	317

a. Model: $log(\frac{\pi(x)}{1-\pi(x)}) = \mu + d_2 + d_3 + d_4 + d_5 + d_6 + e$

```
mdl = glm(cbind(dta$yes, dta$no) ~ factor(department), family = binomial(), data = dta)
summary(mdl)
```

```
Call:
glm(formula = cbind(dta$yes, dta$no) ~ factor(department), family = binomial(),
    data = dta)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4064  -0.4550   0.1456   0.5471   4.1323

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          0.59346    0.06838   8.679   <2e-16 ***
factor(department)2 -0.05059    0.10968  -0.461    0.645
factor(department)3 -1.20915    0.09726 -12.432   <2e-16 ***
factor(department)4 -1.25833    0.10152 -12.395   <2e-16 ***
factor(department)5 -1.68296    0.11733 -14.343   <2e-16 ***
factor(department)6 -3.26911    0.16707 -19.567   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 877.056  on 11  degrees of freedom
Residual deviance:  21.736  on  6  degrees of freedom
AIC: 102.68

Number of Fisher Scoring iterations: 4
```

b. An informal test of fit is deviance/df = 3.6 which would indicate that the model does not fit the data very well. You want deviance/df to be close to 1.

c. Department one is a clear outlier with standardized residuals greater than 4 meaning a lot more females were accepted than what was expected.

d. `residuals(mdl)[1] = -1.4`. Less males than expected were accepted into department 1 by ~1.4 standard deviations. This makes sense because the female acceptance rate is so high and so it probably prevented more males from being accepted.

e. There are many more men than woman in this data and the men overall have a higher overall acceptance rate when you average over department. In 4 of the 6 departments the acceptance rate is higher for females than it is for males which shows conditional on department, femalse have a higher chance of getting accepted.

4

```r
icu = read.csv("icu.csv")
icu = icu[, c("sta", "age", "can", "typ", "ph", "pco", "loc", "sys")]

mdl = glm(sta ~ .+ .*., family = binomial(), data = icu)
summary(mdl)
```

```
Call:
glm(formula = sta ~ . + . * ., family = binomial(), data = icu)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2976  -0.4963  -0.2177   0.0000   2.3761

Coefficients: (2 not defined because of singularities)
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.774e+01  5.646e+04  -0.001   0.9995
age         -1.175e-01  1.252e+03   0.000   0.9999
can         -4.477e+02  1.565e+05  -0.003   0.9977
typ          2.751e+01  5.646e+04   0.000   0.9996
ph           2.238e+00  7.142e+04   0.000   1.0000
pco          2.373e+01  5.763e+04   0.000   0.9997
loc          1.452e+02  5.958e+04   0.002   0.9981
sys          1.881e-01  5.315e+02   0.000   0.9997
age:can      7.408e+00  1.218e+03   0.006   0.9951
age:typ      2.864e-01  1.252e+03   0.000   0.9998
age:ph       2.837e-02  7.820e-02   0.363   0.7168
age:pco     -1.359e-01  1.767e-01  -0.769   0.4418
age:loc     -3.013e+00  5.903e+02  -0.005   0.9959
age:sys     -1.035e-03  6.050e-04  -1.710   0.0872 .
can:typ      1.488e+02  6.754e+04   0.002   0.9982
can:ph       1.411e+02  8.099e+04   0.002   0.9986
can:pco     -3.188e+01  6.214e+04  -0.001   0.9996
can:loc             NA         NA      NA       NA
can:sys     -4.333e-01  4.112e+02  -0.001   0.9992
typ:ph      -3.560e-01  7.142e+04   0.000   1.0000
typ:pco     -1.023e+01  5.763e+04   0.000   0.9999
typ:loc      3.761e+01  5.740e+04   0.001   0.9995
typ:sys     -1.380e-01  5.315e+02   0.000   0.9998
ph:pco       1.267e+00  3.061e+00   0.414   0.6790
ph:loc              NA         NA      NA       NA
ph:sys      -1.814e-02  3.179e-02  -0.571   0.5682
pco:loc     -4.816e+01  8.366e+03  -0.006   0.9954
pco:sys     -5.151e-02  8.609e-02  -0.598   0.5496
loc:sys      6.834e-01  1.287e+02   0.005   0.9958
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 200.16  on 199  degrees of freedom
Residual deviance: 100.01  on 173  degrees of freedom
AIC: 154.01

Number of Fisher Scoring iterations: 21
```

```
mdl.01 = update(mdl   , ~ . - can:loc) ; itr.01 = mdl.01$aic # singularities
mdl.02 = update(mdl.01, ~ . - ph:loc)  ; itr.02 = mdl.02$aic # singularities
mdl.03 = update(mdl.02, ~ . - typ:ph)  ; itr.03 = mdl.03$aic # pvalue = 1.0
mdl.04 = update(mdl.03, ~ . - can:sys) ; itr.04 = mdl.04$aic # pvalue = .999
mdl.05 = update(mdl.04, ~ . - can:pco) ; itr.05 = mdl.05$aic # pvalue = .999
mdl.06 = update(mdl.05, ~ . - typ:sys) ; itr.06 = mdl.06$aic # pvalue = .998
mdl.07 = update(mdl.06, ~ . - age:typ) ; itr.07 = mdl.07$aic # pvalue = .999
mdl.08 = update(mdl.07, ~ . - typ:loc) ; itr.08 = mdl.08$aic # pvalue = .999
mdl.09 = update(mdl.08, ~ . - can:ph)  ; itr.09 = mdl.09$aic # pvalue = .999
mdl.10 = update(mdl.09, ~ . - age:loc) ; itr.10 = mdl.10$aic # pvalue = .995
mdl.11 = update(mdl.10, ~ . - typ:pco) ; itr.11 = mdl.11$aic # pvalue = .996
mdl.12 = update(mdl.11, ~ . - pco:loc) ; itr.12 = mdl.12$aic # pvalue = .980
mdl.13 = update(mdl.12, ~ . - age:ph)  ; itr.13 = mdl.13$aic # pvalue = .897
mdl.14 = update(mdl.13, ~ . - ph:sys)  ; itr.14 = mdl.14$aic # pvalue = .662
mdl.15 = update(mdl.14, ~ . - can:typ) ; itr.15 = mdl.15$aic # pvalue = .562
mdl.16 = update(mdl.15, ~ . - age:pco) ; itr.16 = mdl.16$aic # pvalue = .539
mdl.17 = update(mdl.16, ~ . - age:can) ; itr.17 = mdl.17$aic # pvalue = .096
mdl.18 = update(mdl.17, ~ . - pco:sys) ; itr.18 = mdl.18$aic # pvalue = .060
mdl.19 = update(mdl.18, ~ . - ph:pco)  ; itr.19 = mdl.19$aic # pvalue = .232
mdl.20 = update(mdl.19, ~ . - loc:sys) ; itr.20 = mdl.20$aic # pvalue = .219
```

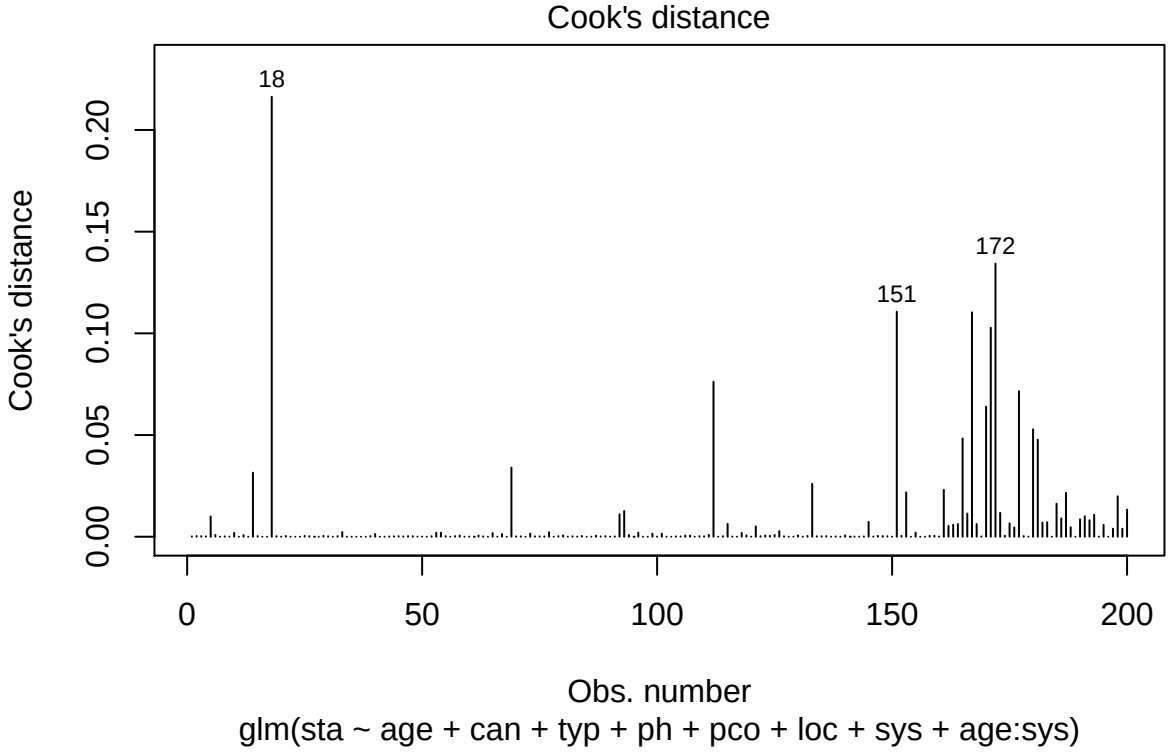
```
summary(mdl.20)
```

```
Call:
glm(formula = sta ~ age + can + typ + ph + pco + loc + sys +
    age:sys, family = binomial(), data = icu)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7440  -0.5469  -0.2961  -0.1231   2.6142

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.642e+01  4.411e+00  -3.721 0.000198 ***
age          2.121e-01  6.421e-02   3.304 0.000954 ***
can          2.413e+00  8.706e-01   2.771 0.005585 **
```

```
typ          2.867e+00  9.328e-01   3.074 0.002115 **
ph           1.859e+00  8.774e-01   2.119 0.034129 *
pco         -2.779e+00  1.063e+00  -2.614 0.008949 **
loc          2.661e+00  6.828e-01   3.897 9.73e-05 ***
sys          7.100e-02  2.956e-02   2.402 0.016293 *
age:sys     -1.287e-03  4.566e-04  -2.818 0.004836 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 200.16  on 199  degrees of freedom
Residual deviance: 128.10  on 191  degrees of freedom
AIC: 146.1

Number of Fisher Scoring iterations: 6


5
```

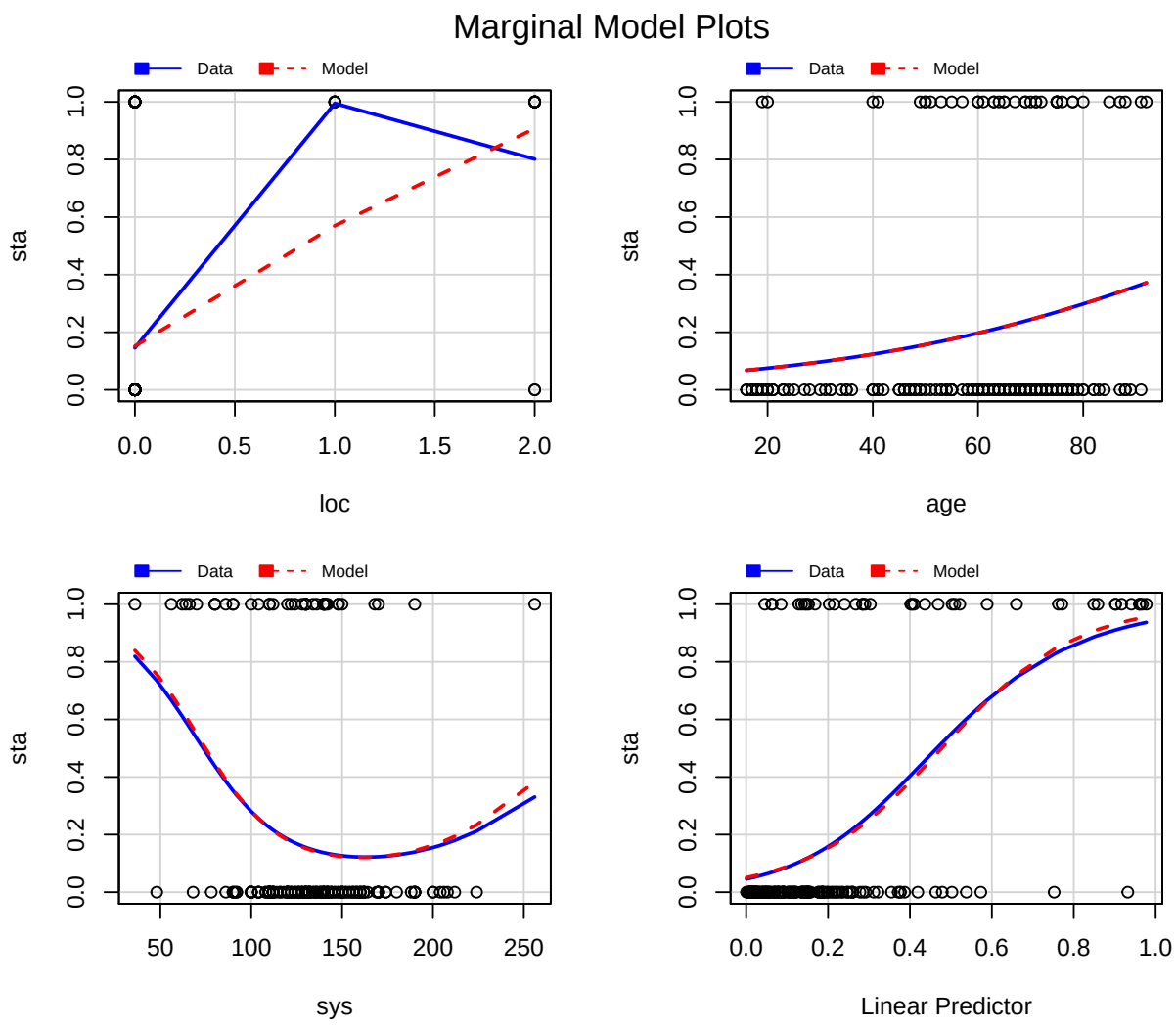
The final model (mdl.20) has the highest residual deviance of all of the models because it has the fewest variables. It only has an AIC near the middle of the pack so it is not considered the best overall model compared to all combinations of main effects and 2 way interactions because it doesnt explain the maximum deviance, however it is the only model where all of the variables are statistically significant. Most of the residual diagnostic plots are not useful when the response variable is 0/1, but you can use the cooks distance plot to assess if any observations have high leverage (far away from the average) and influence. Observations 151, 172, 18 have the highest leverage. Marginal model plots (next page) are also used to assess the fit. Both Sys and Age approximate the data well by the close proximity between the model and data lines. Loc does not appear to model sta very well. The overall model fit (bottom right of marginal model plot) shows a well fitted model.

```
## AIC: Lowest is md.11 (137.9)
c(itr.01, itr.02, itr.03, itr.04, itr.05, itr.06, itr.07, itr.08, itr.09, itr.10,
  itr.11, itr.12, itr.13, itr.14, itr.15, itr.16, itr.17, itr.18, itr.19, itr.20)

 [1] 154.0102 154.0102 152.0102 150.0102 148.0102 146.0102 144.0102
 [8] 142.0102 140.0102 139.9048 137.9546 142.9121 140.9292 139.1432
[15] 143.9143 142.2586 144.8281 146.9906 146.5335 146.1006

## Cooks Distance
cutoff = 4/((200 - length(mdl.20$coefficients)-2))
plot(mdl.20, which=4, cook.levels=cutoff)
```



6

```
Call:  glm(formula = sta ~ factor(can) + factor(typ) + loc + age + sys +
    age:sys, family = binomial(), data = icu)

Coefficients:
 (Intercept)  factor(can)1  factor(typ)1           loc           age
  -14.481110      2.082406      2.791547      1.925665      0.180910
         sys       age:sys
    0.060589     -0.001104

Degrees of Freedom: 199 Total (i.e. Null);  193 Residual
Null Deviance:      200.2
Residual Deviance: 137.2    AIC: 151.2
```



5.25

Lymphocytic.Infiltration	Sex	Osteoblastic.Pathology	Disease.Free.Yes	Disease.Free.No
High	Female	No	3	0
High	Female	Yes	2	0
High	Male	No	4	0
High	Male	Yes	1	0
Low	Female	No	5	0
Low	Female	Yes	3	2
Low	Male	No	5	4
Low	Male	Yes	6	11

a.

```
Call:
glm(formula = cbind(dta$Disease.Free.Yes, dta$Disease.Free.No) ~
    Lymphocytic.Infiltration, family = binomial(), data = dta)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.44956   0.00008   0.00012   0.20659   2.52800

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                    20.06    4357.04   0.005    0.996
Lymphocytic.InfiltrationLow   -19.95    4357.04  -0.005    0.996

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 19.4327  on 7  degrees of freedom
Residual deviance:  8.6256  on 6  degrees of freedom
AIC: 20.671

Number of Fisher Scoring iterations: 18



Call:
glm(formula = cbind(dta$Disease.Free.Yes, dta$Disease.Free.No) ~
    Sex, family = binomial(), data = dta)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4792  -0.1607   0.8416   1.1617   2.3003

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   1.8718     0.7595   2.464   0.0137 *
SexMale      -1.8073     0.8403  -2.151   0.0315 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 19.433  on 7  degrees of freedom
Residual deviance: 13.553  on 6  degrees of freedom
AIC: 25.598

Number of Fisher Scoring iterations: 4
```

```
Call:
glm(formula = cbind(dta$Disease.Free.Yes, dta$Disease.Free.No) ~
    Osteoblastic.Pathology, family = binomial(), data = dta)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7360   0.1389   1.1688   1.3385   1.7134

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)                1.4469     0.5557   2.604  0.00922 **
Osteoblastic.PathologyYes -1.5270     0.6849  -2.230  0.02578 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 19.433  on 7  degrees of freedom
Residual deviance: 13.898  on 6  degrees of freedom
AIC: 25.943

Number of Fisher Scoring iterations: 4
```

b. Its infinite because the Lymphocytic Infiltration is high all of the counts are disease free.

```
Call:
glm(formula = cbind(dta$Disease.Free.Yes, dta$Disease.Free.No) ~
    Lymphocytic.Infiltration + Sex + Osteoblastic.Pathology,
    family = binomial(), data = dta)

Deviance Residuals:
       1         2         3         4         5         6         7
 0.00002   0.00003   0.00005   0.00005   1.07088  -0.51727  -0.36813
       8
 0.27912

Coefficients:
                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                   23.4920 11084.3781   0.002   0.9983
Lymphocytic.InfiltrationLow  -21.3842 11084.3781  -0.002   0.9985
SexMale                       -1.6362     0.9123  -1.794   0.0729 .
Osteoblastic.PathologyYes     -1.2204     0.7712  -1.582   0.1135
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 19.4327  on 7  degrees of freedom
Residual deviance:  1.6278  on 4  degrees of freedom
AIC: 17.673

Number of Fisher Scoring iterations: 20
```

c. $.002 < 1.96 = qnorm(.975), (1 - pnorm(.002)) * 2 = .998$

d. 95% Confidence interval: (-Inf, 991.3)

```
confint(mdl)
```

```
                                  2.5 %       97.5 %
(Intercept)                 -822.834757           NA
Lymphocytic.InfiltrationLow          NA 991.28251566
SexMale                      -3.699935   0.02499745
Osteoblastic.PathologyYes    -2.827264   0.24902144
```