

STATS 4F03/6F03 Test #2 SOLUTIONS

Problems 2.25, 2.26

I have computed the exact-test P-values using the hypergeometric distribution and using the `fisher.test()` function. For the hypergeometric, I took the top right cell as the random variable, making the one-sided test a left-tail test. This seemed to be more convenient.

```
> for(x in 0:5) print(matrix(c(23-x,13+x,x,5-x),nrow=2))
      [,1] [,2]
[1,]   23   0
[2,]   13   5
      [,1] [,2]
[1,]   22   1
[2,]   14   4
      [,1] [,2]
[1,]   21   2
[2,]   15   3
      [,1] [,2]
[1,]   20   3
[2,]   16   2
      [,1] [,2]
[1,]   19   4
[2,]   17   1
      [,1] [,2]
[1,]   18   5
[2,]   18   0

> dhyper(0:5,23,18,5)
[1] 0.01143318 0.09391538 0.27548512 0.36157422 0.21269072 0.04490137
```

One-sided P-value:

```
> phyper(2,23,18,5)
[1] 0.3808337
```

One-sided mid-P:

```
> phyper(1,23,18,5)+0.5*dhyper(2,23,18,5)
[1] 0.2430911
```

Conclusion from the one-sided test: There is no evidence from these data ($P = 0.38$, mid-P = 0.24) that the odds of controlling cancer are greater with surgery than with radiation treatment.

Two-sided P-value:

```
> sum(dhyper(0:5,23,18,5)[dhyper(0:5,23,18,5) <= dhyper(2,23,18,5)])
[1] 0.6384258
```

Two-sided mid-P:

```
> sum(dhyper(0:5,23,18,5)[dhyper(0:5,23,18,5) <
dhyper(2,23,18,5)])+0.5*dhyper(2,23,18,5)
[1] 0.5006832
```

Conclusion from the two-sided test: There is no evidence from these data ($P = 0.64$, mid-P = 0.50) that the odds of controlling cancer are different with surgery than with radiation treatment.

The smallest possible P-value would result when $X = 0$ (the first table in the list of tables above). Because this is the smallest possible hypergeometric probability, the P-value would be the same for a two-sided test as for a one-sided test (left-tail in my case). Note that a 1% test could never reject the hypothesis.

```
> dhyper(0,23,18,5)
```

```
[1] 0.01143318
```

Since $X = 0$ is the extreme tail observation, the smallest possible mid-P is half of the smallest possible P. Note that using mid-P a 1% test could reject the hypothesis but a 0.5% test could never reject.

```
> 0.5*dhyper(0,23,18,5)
```

```
[1] 0.005716589
```

Here is the sample odds ratio and a confidence interval for the true odds ratio computed using Woolf's formula:

```
> (21*3)/(15*2)
```

```
[1] 2.1
```

```
>> ((21*3)/(15*2))*exp(c(-1,1)*qnorm(0.975)*sqrt((1/21)+(1/15)+(1/2)+(1/3)))
```

```
[1] 0.3116097 14.1523194
```

fisher.test() gives the same P-values as I got with the hypergeometric but the estimated odds ratio and confidence intervals are different because they are true conditional ML estimates based on the non-central hypergeometric.

```
> fisher.test(matrix(c(21,15,2,3),nrow=2))
```

```
Fisher's Exact Test for Count Data
```

```
data: matrix(c(21, 15, 2, 3), nrow = 2)
```

```
p-value = 0.6384
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
0.2089115 27.5538747
```

```
sample estimates:
```

```
odds ratio
```

```
2.061731
```

```
> fisher.test(matrix(c(21,15,2,3),nrow=2), alt="greater")
```

```
Fisher's Exact Test for Count Data
```

```
data: matrix(c(21, 15, 2, 3), nrow = 2)
```

```
p-value = 0.3808
```

```
alternative hypothesis: true odds ratio is greater than 1
```

```
95 percent confidence interval:
```

```
0.2864828 Inf
```

```
sample estimates:
```

```
odds ratio
```

```
2.061731
```

Here for comparison is the two-sided P-value computed by Pearson's chi-square with and without continuity correction.

```
> chisq.test(matrix(c(21,15,2,3),nrow=2))
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: matrix(c(21, 15, 2, 3), nrow = 2)
```

```
X-squared = 0.086, df = 1, p-value = 0.7694
```

```
Warning message:
```

```
Chi-squared approximation may be incorrect in: chisq.test(matrix(c(21, 15, 2, 3), nrow = 2))
```

```
> chisq.test(matrix(c(21,15,2,3),nrow=2),correct=F)
```

```
Pearson's Chi-squared test
```

```
data: matrix(c(21, 15, 2, 3), nrow = 2)
```

```
X-squared = 0.5992, df = 1, p-value = 0.4389
```

```
Warning message:
```

```
Chi-squared approximation may be incorrect in: chisq.test(matrix(c(21, 15, 2,
3), nrow = 2), correct = F)
```

Finally, here is a logistic analysis; the P-value is, of course, very close to that obtained with Pearson's Chi-square without continuity correction as the model and the estimates are the same but the glm uses G^2 instead of X^2 to test the fit.

```
> q2.25 <- data.frame(cc=c(21,15),cnc=c(2,3),treat=c("Su","RT"))
> anova(glm(cbind(cc,cnc)~treat, q2.25, family=binomial(link=logit)),
test="Chis")
```

```
Analysis of Deviance Table
```

```
Model: binomial, link: logit
```

```
Response: cbind(cc, cnc)
```

```
Terms added sequentially (first to last)
```

	Df	Deviance	Resid. Df	Resid. Dev	P(> Chi)
NULL			1	0.59476	
treat	1	0.59476	0	6.661e-16	0.44058

```
> summary(glm(cbind(cc,cnc)~treat, q2.25, family=binomial(link=logit)))
```

```
Call:
```

```
glm(formula = cbind(cc, cnc) ~ treat, family = binomial(link = logit),
data = q2.25)
```

```
Deviance Residuals:
```

```
[1] 0 0
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.6094	0.6325	2.545	0.0109 *
treatSu	0.7419	0.9735	0.762	0.4460

```
---
```

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

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

```
Null deviance: 5.9476e-01 on 1 degrees of freedom
```

```
Residual deviance: 6.6613e-16 on 0 degrees of freedom
```

```
AIC: 9.3348
```

```
Number of Fisher Scoring iterations: 4
```

The 95% confidence interval for the odds ratio calculated from the logistic analysis agrees with the one calculated by Woolf's formula; I re-entered the estimated log odds ratio and standard deviation from the printout above so the result is only accurate to 4 significant digits.

```
> exp(.7419)
```

```
[1] 2.099922
```

```
> exp(.7419+c(-1,1)*1.960*.9736)
```

```
[1] 0.31150 14.15625
```

Problem 3.15

I have done the analysis first with cross-tabulated data and the `mantelhaen.test()` and `breslowday.test()` functions, then with the data frame and logistic analysis.

```
> merit
  freq district pay race
1    24      NC  Y  B
2    10      NE  Y  B
3     5      NW  Y  B
4    16      SE  Y  B
5     7      SW  Y  B
6     9      NC  N  B
7     3      NE  N  B
8     4      NW  N  B
9     7      SE  N  B
10    4      SW  N  B
11   47      NC  Y  W
12   45      NE  Y  W
13   57      NW  Y  W
14   54      SE  Y  W
15   59      SW  Y  W
16   12      NC  N  W
17    8      NE  N  W
18    9      NW  N  W
19   10      SE  N  W
20   12      SW  N  W
> meritx <- xtabs(freq~race+pay+district, merit)
> meritx
, , district = NC

      pay
race  N  Y
  B   9 24
  W  12 47

, , district = NE

      pay
race  N  Y
  B   3 10
  W   8 45

, , district = NW

      pay
race  N  Y
  B   4  5
  W   9 57

, , district = SE

      pay
race  N  Y
  B   7 16
  W  10 54

, , district = SW
```

```

      pay
race  N  Y
   B   4  7
   W  12 59

```

With the exception of district NW, which has a P-value close to 5%, the partial tables do not show association between race and merit pay.

```
> apply(meritx, 3, fisher.test)
```

```
$NC
```

```
Fisher's Exact Test for Count Data
```

```

data:  array(newX[, i], d.call, dn.call)
p-value = 0.4509
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.4728644 4.4115657
sample estimates:
odds ratio
 1.462374

```

```
$NE
```

```
Fisher's Exact Test for Count Data
```

```

data:  array(newX[, i], d.call, dn.call)
p-value = 0.6782
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.243233 8.758798
sample estimates:
odds ratio
 1.672719

```

```
$NW
```

```
Fisher's Exact Test for Count Data
```

```

data:  array(newX[, i], d.call, dn.call)
p-value = 0.0431
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.8170578 28.1676045
sample estimates:
odds ratio
 4.912599

```

```
$SE
```

```
Fisher's Exact Test for Count Data
```

```

data:  array(newX[, i], d.call, dn.call)
p-value = 0.1368
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.6448157 8.1625569

```

```
sample estimates:
odds ratio
  2.336334
```

```
$SW
```

```
Fisher's Exact Test for Count Data
```

```
data: array(newX[, i], d.call, dn.call)
p-value = 0.212
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.5120718 13.1045517
sample estimates:
odds ratio
  2.765975
```

```
> apply(meritx,3,chisq.test)
$NC
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: array(newX[, i], d.call, dn.call)
X-squared = 0.251, df = 1, p-value = 0.6164
```

```
$NE
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: array(newX[, i], d.call, dn.call)
X-squared = 0.0766, df = 1, p-value = 0.7819
```

```
$NW
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: array(newX[, i], d.call, dn.call)
X-squared = 3.3164, df = 1, p-value = 0.06859
```

```
$SE
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: array(newX[, i], d.call, dn.call)
X-squared = 1.5124, df = 1, p-value = 0.2188
```

```
$SW
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data: array(newX[, i], d.call, dn.call)
X-squared = 1.225, df = 1, p-value = 0.2684
```

The Mantel-Haenszel test shows that if there is a common odds ratio for all 5 districts, it is significantly different from 1 ($P = 0.008$ with continuity correction, $P = 0.005$ without), hence there is an association between race and merit pay. The odds that a white will get merit pay are more than twice the odds that a black will get merit pay.

```
> mantelhaen.test(meritx)
```

Mantel-Haenszel chi-squared test with continuity correction

```
data: meritx
Mantel-Haenszel X-squared = 6.9797, df = 1, p-value = 0.008244
alternative hypothesis: true common odds ratio is not equal to 1
95 percent confidence interval:
 1.244268 3.769779
sample estimates:
common odds ratio
      2.165783
```

```
> mantelhaen.test(meritx, correct=F)
```

Mantel-Haenszel chi-squared test without continuity correction

```
data: meritx
Mantel-Haenszel X-squared = 7.8149, df = 1, p-value = 0.005182
alternative hypothesis: true common odds ratio is not equal to 1
95 percent confidence interval:
 1.244268 3.769779
sample estimates:
common odds ratio
      2.165783
```

The exact Mantel-Haenszel test gives a result that is not much different; the P-value is a bit smaller and the confidence interval is a bit wider.

```
> mantelhaen.test(meritx, correct=F, exact=T)
```

Exact conditional test of independence in 2 x 2 x k tables

```
data: meritx
S = 27, p-value = 0.008428
alternative hypothesis: true common odds ratio is not equal to 1
95 percent confidence interval:
 1.197300 3.933736
sample estimates:
common odds ratio
      2.182267
```

The Breslow-Day test shows that the odds ratio does not differ significantly between districts ($P = 0.71$), whether the common odds ratio is the marginal odds ratio (default in the Breslow-Day test) or the Mantel-Haenszel estimator.

```
> breslowday.test(meritx)
```

```
      OR      Stat      df  pvalue
2.237192 2.137252 4.000000 0.710532
```

```
> breslowday.test(meritx, mantelhaen.test(meritx, correct=F)$estimate)
```

```
OR.common odds ratio      Stat      df
      2.1657826      2.1506942      4.0000000
      pvalue
      0.7080655
```

In the logistic analysis, the test of race:district interaction ($P = 0.72$) is equivalent to the Breslow-Day test and the test of race main effect ($P = 0.007$) is equivalent to the Mantel-Haenszel test.

```
> meritl
```

```

      yes no district race
1     24  9         NC   B
2     10  3         NE   B
3      5  4         NW   B
4     16  7         SE   B
5      7  4         SW   B
11    47 12         NC   W
12    45  8         NE   W
13    57  9         NW   W
14    54 10         SE   W
15    59 12         SW   W

```

```
> anova(glm(cbind(yes,no)~district*race, meritl, family=binomial(link=logit)),
test="Chis")
```

```
Analysis of Deviance Table
```

```
Model: binomial, link: logit
```

```
Response: cbind(yes, no)
```

```
Terms added sequentially (first to last)
```

	Df	Deviance	Resid. Df	Resid. Dev	P(> Chi)
NULL			9	10.6649	
district	4	1.2024	5	9.4624	0.8777
race	1	7.3915	4	2.0710	0.0066
district:race	4	2.0710	0	-1.266e-14	0.7227

```
> summary(glm(cbind(yes,no)~district+race, meritl,
family=binomial(link=logit)))
```

```
Call:
```

```
glm(formula = cbind(yes, no) ~ district + race, family = binomial(link =
logit),
    data = meritl)
```

```
Deviance Residuals:
```

1	2	3	4	5	11	12	13
0.60191	0.30311	-0.97042	-0.09608	-0.30707	-0.53319	-0.18583	0.47422
14	15						
0.07216	0.15054						

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.74947	0.29581	2.534	0.01129 *
districtNE	0.25837	0.42067	0.614	0.53909
districtNW	0.13836	0.40517	0.341	0.73273
districtSE	0.12087	0.37287	0.324	0.74581
districtSW	0.00445	0.38486	0.012	0.99077
raceW	0.79129	0.28532	2.773	0.00555 **

```
---
```

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

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

```

Null deviance: 10.665  on 9  degrees of freedom
Residual deviance:  2.071  on 4  degrees of freedom
AIC: 49.437

```

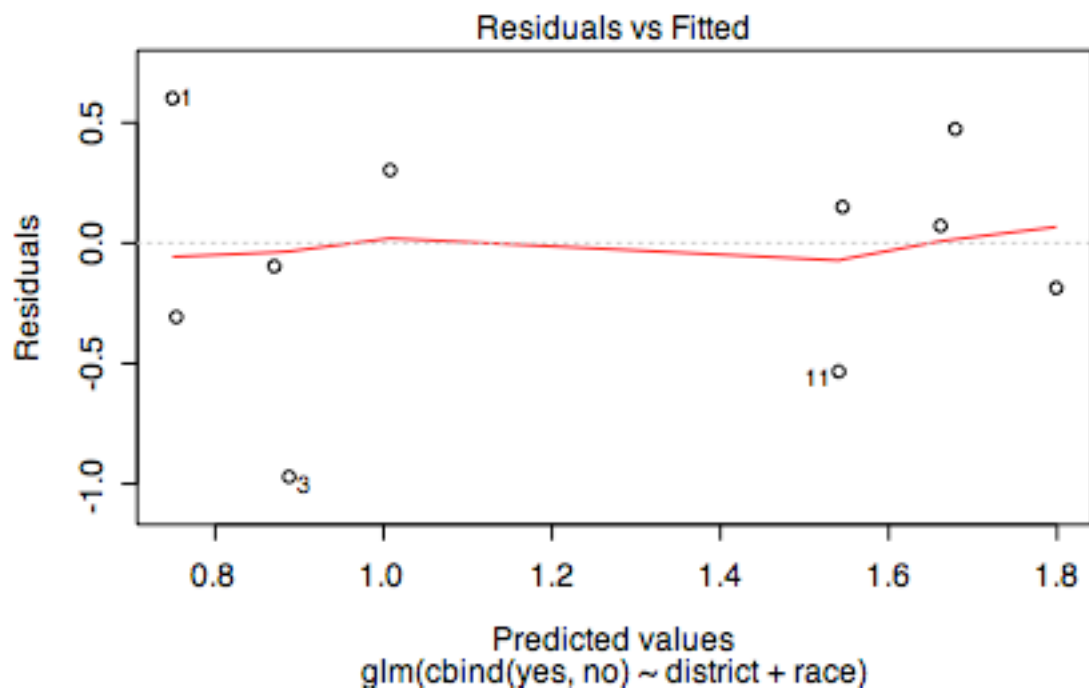

Number of Fisher Scoring iterations: 4

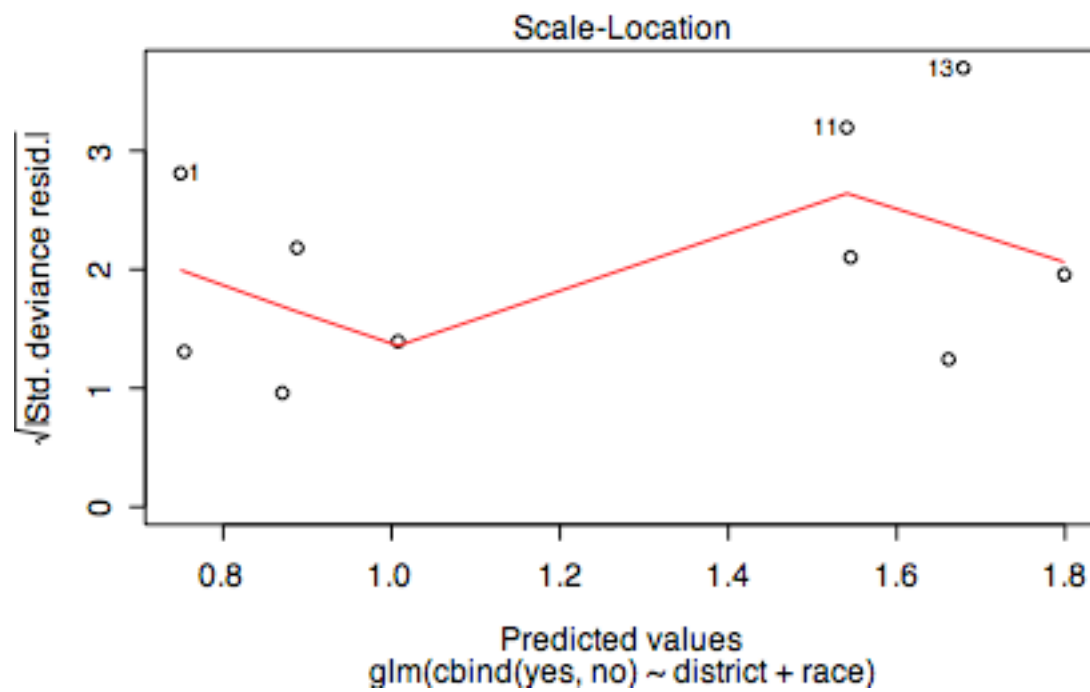
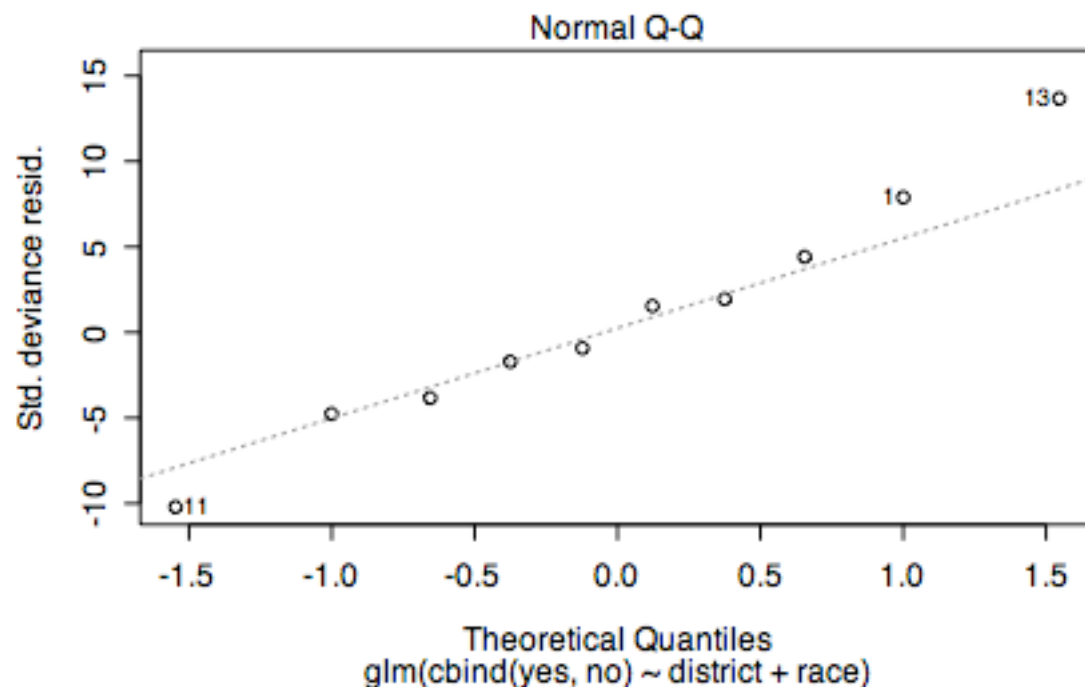
The regression coefficient β and its standard error give a confidence interval for the common odds ratio.

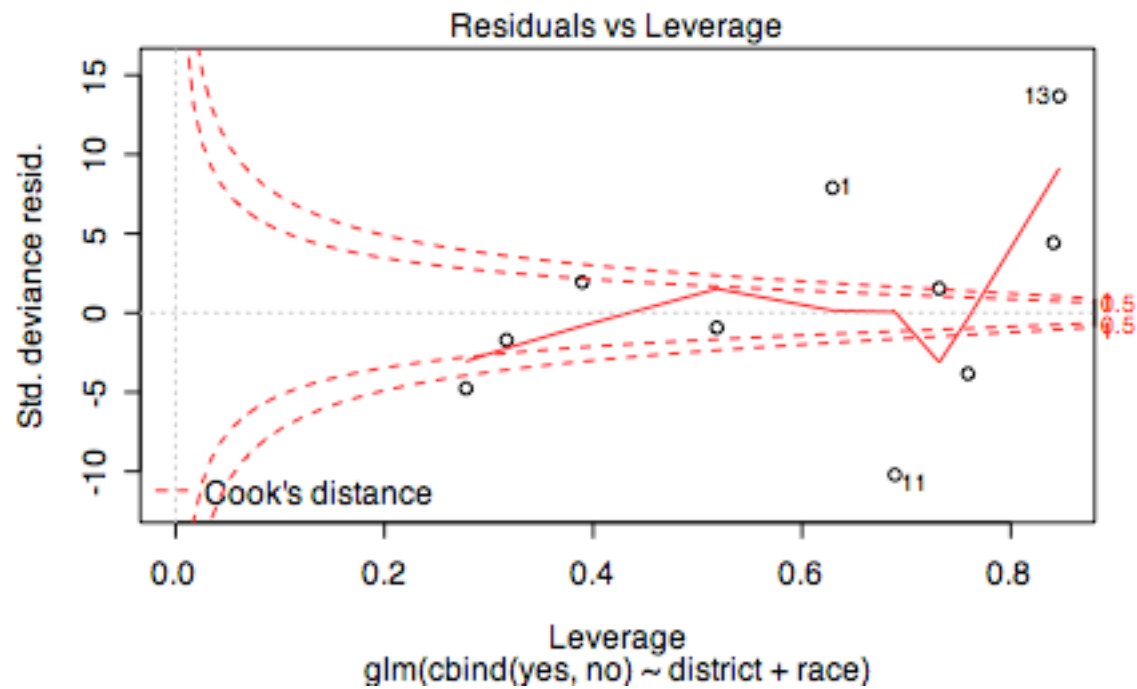
```
> exp(.79129)
[1] 2.206241
> exp(.79129 + c(-1,1)*qnorm(.975)*0.28532)
[1] 1.261212 3.859381
```

The following command gives 4 diagnostic plots.

```
> plot(glm(cbind(yes,no)~district+race, merit1, family=binomial(link=logit)))
```







The diagnostic plots show observations 1 (NC black), 11 (NC white), 3 (NW black) and 13 (NW white) as being influential (large residual and high leverage). To interpret the plots, remember that there are 10 points but only 4 degrees of freedom. The plot of residual versus predicted, for example, shows the 5 points for blacks on the left (merit pay less likely) and the 5 points for whites on the right (more likely) and the two clusters of points are nearly top-to-bottom flips of each other.