Chong Woon Kiat A0209349X - ST5213 Assignment 2

Task 1

(a)
Under H_o: medical aid status and participation in interview are independent, the observed (and expected in bracket) count is given by:

	No interview	Interview	Total
Had medical aid	195 (177.9)	46 (63.1)	241
No medical aid	979 (996.1)	370 (352.9)	1349
Total	1174	416	1590

The Pearson $X^2 = \sum_{ij} \frac{(n_{ij}-\mu_{ij})^2}{\mu_{ij}} = 7.36$

X^2 follows χ_1^2 distribution approximately, the p-value = $P(\chi_1^2 > 7.36) = 0.0067$. Hence, the null hypothesis that they are independent is rejected. There is strong evidence of an association between medical aid status and participation in interviews.

The odds ratio between medical aid status and participation in interviews is given by
$\theta = \frac{n_{11}n_{22}}{n_{21}n_{12}} = \frac{195 \times 370}{979 \times 46} = 1.60$
The odds of a child with medical aid not participating in the interview is 1.60 times the odds of a child with no medical aid not participating in the interview.

(b)

White	No interview	Interview	Total
Had medical aid	x	114-x	114
No medical aid	126-x	x-102	24
Total	126	12	138

The range of x is [102,114].

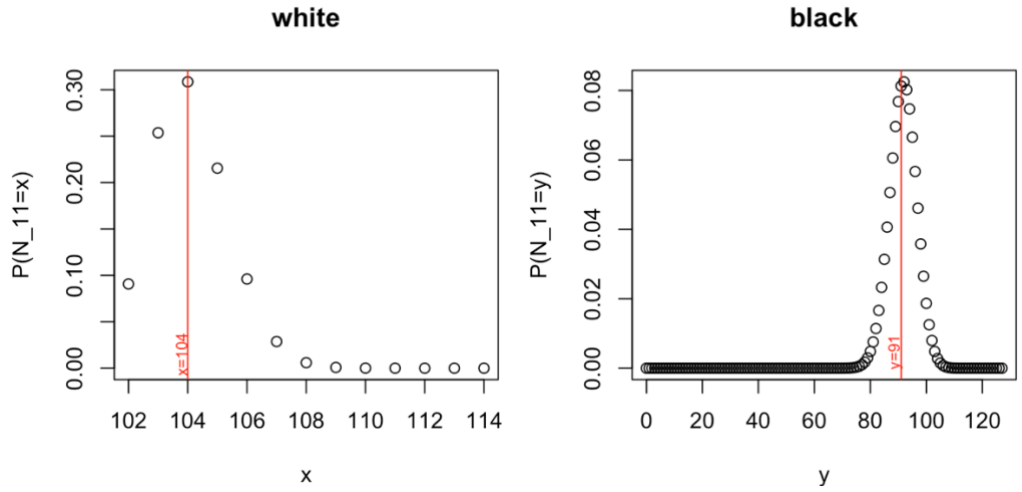
Black	No interview	Interview	Total
Had medical aid	y	127-y	127
No medical aid	1048-y	y+277	1325
Total	1048	404	1452

The range of y is [0,127].

To test for $H_0 : \theta = 1$ and the two sided alternative $H_1 : \theta \neq 1$, the p-value is the two-tailed sum of probabilities of tables no more likely than the observed table, where the probability of table follows hypergeometric distribution:

For white children is $P(N_{11} = x) = \frac{\binom{114}{x}\binom{24}{126-x}}{\binom{138}{126}}$, $x \in [102,114]$.

For black children is $P(N_{11} = y) = \frac{\binom{127}{y}\binom{1325}{1048-y}}{\binom{1452}{1048}}, y \in [0,127]$.



- For white children, no other table has probability higher than the observed table of $P(N_{11} = 104) = \frac{\binom{114}{104}\binom{24}{22}}{\binom{138}{126}} = 0.3085$. Hence, the p-value = 1.
- For black children, the observed table has probability of $P(N_{11} = 91) = \frac{\binom{127}{91}\binom{1325}{957}}{\binom{1452}{1048}} = 0.08129$. Only when N_{11}=92, the probability of table $P(N_{11} = 92) = \frac{\binom{127}{92}\binom{1325}{956}}{\binom{1452}{1048}} = 0.0825$, is higher than $P(N_{11} = 91)$. Hence, the p-value = 1-0.0825 = 0.9175.
- The p-value for both test >0.05 and therefore hypothesis that θ = 1 is not rejected at 5% significance level. There is insignificant evidence of an association between medical aid status and participation in the interview for both white and black children.

(c)
Simpson's paradox happens as there is strong association between the controlled variable (the children's race) and the response variables (medical aid status and interview participation).

	No interview	Interview
White	126	12
Black	1048	404

	Had medical aid	No medical aid
White	114	24
Black	127	1325

If the table is collapsed over medical aid status (yielding table on the left), the odds ratio is $\frac{126\times404}{1048\times12} = 4.05 > 1$, suggesting that the odds of a white child not participating in the interview is 4.05 times the odds of a black child.

Also, if the table is collapsed over interview participation (yielding table on the right), the odds ratio is $\frac{114\times1325}{127\times24} = 49.56 > 1$, suggesting that the odds of a white child having medical aid is 49.56 times that odds of a black child.

The marginal odds of not participating in the interview appear to be higher for children who had medical aid because white children are more likely to receive medical aid but have lower tendency to attend an interview.

Task 2

Smoking, Family structure, Race, Gender, Age are denoted as S, F, R, G, A.

(a)
(i) The minimal model is (S, F, RGA).

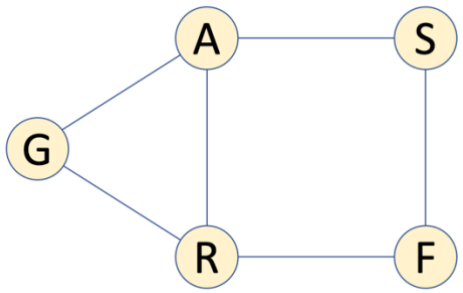
(ii) A loglinear model is built by first fitting the full model and applying drop-in-deviance test to drop terms which are not significant one at a time.

The preferred model is (FS, AS, FR, RGA). The table on the right gives the coefficient of the fitted model.

Since deviance follows χ_{df} approximately, the model has a deviance of 21.1 on 19 degree of freedom (p-value = 0.33), hence the hypothesis that the models are adequate is not rejected at 5% confidence level.

Parameter	Est.	Std.Err
(Intercept)	3.19152	0.16334
FMotherOnly	-0.26257	0.17112
RWhite	2.76143	0.1696
GFemale	0.06062	0.2011
A2	-0.58812	0.23921
SSome	-2.43706	0.1159
FMotherOnly:RWhite	**-1.71321**	**0.18761**
RWhite:GFemale	0.03658	0.21075
RWhite:A2	-0.38009	0.25283
GFemale:A2	-0.68478	0.37792
FMotherOnly:SSome	**0.55196**	**0.21095**
A2:SSome	**0.39651**	**0.19148**
RWhite:GFemale:A2	0.14457	0.40032

(iii) The association graph on the right is a graphical model as all interraction corresponding to the cliques are included as sufficient marginals (FS, AS, FR, RGA). An example of a non-graphical model that has the same association graph is (FS, AS, FR, AG, GR, RA).



(iv)
- S and F are conditionally independent of G given A and R, hence the fitted model can be collapsed over G.
- S is conditionally independent of G and R, given A and F ($\theta_{SG} = \theta_{SR} = 1$), hence the fitted model can be collapsed over {G, R}.
- F is conditionally independent of A and G, given S and R ($\theta_{FA} = \theta_{FG} = 1$), hence the fitted model can be collapsed over {A, G}.

- However, it is unable to collapse over {A, G, R} to study the S and F association, over {G, R, F} to study the S and A association, and over {G, A, S} to study the R and F association.

- Both response variables S and F do not occur in three-factor term, the **conditional odds ratio** between S or F and each variable is the same at each combination of levels of the other variables:
  - $\theta_{FS} = \exp(0.552) = 1.737$: The odds of not smoking for a seventh grader with both parents is 1.737 times that of seventh grader with mother only.
  - $\theta_{AS} = \exp(0.397) = 1.487$: The odds of not smoking for a seventh grader of age 12 or younger is 1.487 times that of age 13 or older.
  - $\theta_{FR} = \exp(-1.713) = 0.180$: The odds of a black seventh grader having both parents is 0.180 times that of a white seventh grader.

(v) The zero cell in the table is a random zero, hence its $\mu > 0$ and does contribute to likelihood. Also, none of the sufficient marginals of FS, AS, FR or RGA equals zero. Hence, it does not affect analysis.

(b)
(i) The minimal model is (S, FRGA).

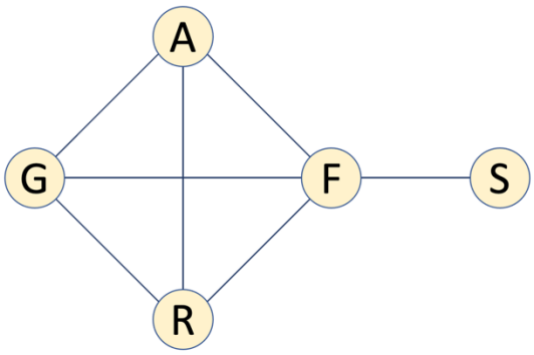
(ii) A loglinear model is built by first fitting the full model and applying drop-in-deviance test to drop terms which are not significant one at a time.

The preferred model is (FS, FRGA). The table on the right gives the coefficient of the fitted model.

Since deviance follows χ_{df} approximately, the model has a deviance of 21.9 on 14 degree of freedom (p-value = 0.08), hence the hypothesis that the models are adequate is not rejected at 5% confidence level.

Parameter	Est.	Std.Err
(Intercept)	3.274	0.186
FMotherOnly	-0.487	0.297
RWhite	2.687	0.192
GFemale	-0.071	0.267
A2	-0.728	0.325
SSome	-2.330	0.100
FMotherOnly:RWhite	-1.643	0.329
FMotherOnly:GFemale	0.305	0.407
RWhite:GFemale	0.146	0.276
FMotherOnly:A2	0.423	0.480
RWhite:A2	-0.245	0.338
GFemale:A2	-0.488	0.518
FMotherOnly:SSome	**0.558**	**0.211**
FMotherOnly:RWhite:GFemale	-0.120	0.451
FMotherOnly:RWhite:A2	-0.072	0.540
FMotherOnly:GFemale:A2	-0.439	0.759
RWhite:GFemale:A2	-0.014	0.537
FMotherOnly:RWhite:GFemale:A2	0.147	0.850

(iii) The association graph on the right is a graphical model as all interaction corresponding to the cliques are included as sufficient marginals (FS, FRGA). An example of a non-graphical model that has the same association graph is (FS, AR, GF, AG, GR, RF, FA).



(iv)
S is conditionally independent of {A, G, R} given F ($\theta_{SA} = \theta_{SG} = \theta_{SR} = 1$).
Hence the fitted model can be collapsed over {A, G, R} to study the association of F and S:

Marginal Count	No Smoking	Some Smoking
Both Parents	1120	109
Mother Only	200	34

With the table collapsed over {A, G, R}, the **marginal odds** of not smoking for a seventh grader with both parents is $\theta_{SF} = \frac{1120 \times 34}{200 \times 109} = 1.747$ times that of seventh grader with mother only, which is the same as the **conditional odds** exp(0.558) = 1.747 at any combination of levels of the other variables as the model can be collapsed over {A, G, R}.

(v) The zero cell in the table is a random zero, hence it $\mu > 0$ and does contribute to likelihood. Also, none of the sufficient marginals of FS, FRGA equals zero. Hence, it does not affect analysis.

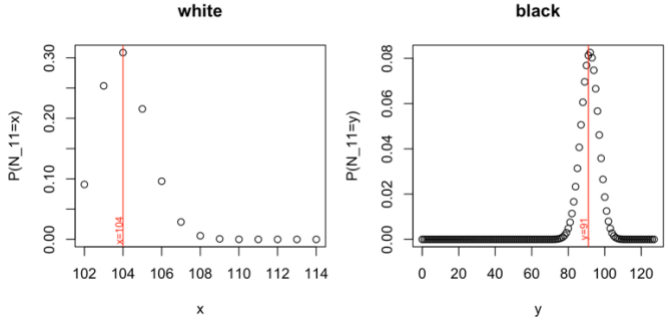
(vi) the logit model equivalent to the above loglinear model is logit(S) ~ F:

logit[P(S=Some | A = a, G = g, R = r, F = f)] = $\beta_0 + \beta_f F$
where
$\beta_f = \lambda_{22}^{FS} = 0.558$
$\beta_0 = -2.330$

Appendix

Task 1

```
> nij <- cbind(c(195,979),c(46,370))
>
> (muhat <- outer(rowSums(nij), colSums(nij), "*")/sum(nij))
         [,1]     [,2]
[1,] 177.9459  63.05409
[2,] 996.0541 352.94591
>
> (X2 <- sum((nij - muhat)^2/muhat))
[1] 7.363053
>
> pchisq(X2, 1, lower.tail=FALSE)
[1] 0.006657763
>
> (odd <- nij[1,1]*nij[2,2]/(nij[1,2]*nij[2,1]))
[1] 1.602123
>
> par(mfrow=c(1,2),mar=c(4,4,4,1))
>
> dhyper(104,114,24,126) #probability of observed table for white
[1] 0.3085378
> plist <- dhyper(102:114,114,24,126)
> i <- dhyper(102:114,114,24,126) <= dhyper(104,114,24,126)
> sum(plist[i]) #p-value
[1] 1
>
> plot(102:114, plist, main="white", xlab = "x", ylab = "P(N_11=x)")
> abline(v=104, col="red")
> text(104-0.2, 0.02, "x=104  ", col = "red",srt=90, cex=0.7)
>
>
> dhyper(91,127,1325,1048) #probability of observed table for black
[1] 0.08128866
> plist <- dhyper(0:127,127,1325,1048)
> i <- dhyper(0:127,127,1325,1048) <= dhyper(91,127,1325,1048)
> sum(plist[i]) #p-value
[1] 0.9175045
>
> plot(0:127, plist, main="black", xlab = "y", ylab = "P(N_11=y)")
> abline(v=91, col="red")
> text(91-2, 0.006, "y=91  ", col = "red",srt=90, cex=0.7)
```



Task 2

```
> F <- c("BothParents", "MotherOnly")
> R <- c("Black","White")
> G <- c("Male","Female")
> A <- c("1","2")
> S <- c("None","Some")
>
> dat<-expand.grid(S=S,A=A,G=G,R=R,F=F)
> dat$Count <- c(27,2,12,2,23,4,7,1,394,32,142,19,421,38,94,11,18,1,13,1,24,0,4,3,48,6,25,4,55,15,13,4)
> ftable(xtabs(Count~F+R+G+A+S,dat))
                    S None Some
F           R     G      A
BothParents Black Male   1     27    2
                         2     12    2
                  Female 1     23    4
                         2      7    1
            White Male   1    394   32
                         2    142   19
                  Female 1    421   38
                         2     94   11
MotherOnly  Black Male   1     18    1
                         2     13    1
                  Female 1     24    0
                         2      4    3
            White Male   1     48    6
                         2     25    4
                  Female 1     55   15
                         2     13    4
```

```
#################### (a) ########################

> fm <- glm(Count ~ F*R*G*A*S, dat, family = poisson)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F * R * G * A * S
        Df Deviance    AIC    LRT Pr(>Chi)
<none>       0.0000 201.68
F:R:G:A:S  1   3.6389 203.32 3.6389  0.05644 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:G:S +
    F:R:A:S + F:G:A:S + R:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       3.6389 203.32
F:R:G:A  1   3.7082 201.39 0.06931   0.7923
F:R:G:S  1   3.8445 201.53 0.20561   0.6502
F:R:A:S  1   6.2708 203.95 2.63190   0.1047
F:G:A:S  1   4.2643 201.95 0.62544   0.4290
R:G:A:S  1   3.9912 201.67 0.35232   0.5528
>
> fm<- update(fm, .~. - A:G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:G:S + F:R:A:S +
    F:G:A:S + R:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       3.7082 201.39
F:R:G:S  1   3.9883 199.67 0.28011   0.5966
F:R:A:S  1   6.2722 201.95 2.56399   0.1093
F:G:A:S  1   4.3730 200.05 0.66477   0.4149
R:G:A:S  1   4.0869 199.77 0.37868   0.5383
>
> fm<- update(fm, .~. - S:G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:A:S + F:G:A:S +
```

```
   R:G:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       3.9883 199.67
F:R:G    1   4.0933 197.78 0.10501   0.7459
F:R:A:S  1   6.6855 200.37 2.69714   0.1005
F:G:A:S  1   4.5023 198.18 0.51400   0.4734
R:G:A:S  1   4.1993 197.88 0.21096   0.6460
>
> fm<- update(fm, .~. - G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + G:S + A:S + F:R:A + F:G:A + R:G:A + F:R:S + F:G:S +
   R:G:S + F:A:S + R:A:S + G:A:S + F:R:A:S + F:G:A:S + R:G:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       4.0933 197.78
F:R:A:S  1   6.8337 198.51 2.74031  0.09785 .
F:G:A:S  1   4.6397 196.32 0.54639  0.45980
R:G:A:S  1   4.3421 196.02 0.24872  0.61798
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:R)
> drop1(fm, test="Chisq") #note AGR cannot be dropped
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + G:S + A:S + F:R:A + F:G:A + R:G:A + F:R:S + F:G:S +
   R:G:S + F:A:S + R:A:S + G:A:S + F:R:A:S + F:G:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       4.3421 196.02
R:G:A    1   4.4092 194.09 0.06718  0.79549
R:G:S    1   4.6572 194.34 0.31515  0.57454
F:R:A:S  1   7.3493 197.03 3.00726  0.08289 .
F:G:A:S  1   5.0310 194.71 0.68892  0.40653
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + G:S + A:S + F:R:A + F:G:A + R:G:A + F:R:S + F:G:S +
   F:A:S + R:A:S + G:A:S + F:R:A:S + F:G:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       4.6572 194.34
R:G:A    1   4.6919 192.37 0.03469  0.85225
F:R:A:S  1   7.7118 195.39 3.05456  0.08051 .
F:G:A:S  1   5.5376 193.22 0.88042  0.34809
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:A + F:G:A + R:G:A + F:R:S + F:G:S +
    F:A:S + R:A:S + G:A:S + F:R:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       5.5376 193.22
F:G:A    1   6.6038 192.28 1.06614  0.30182
R:G:A    1   5.5821 191.26 0.04448  0.83297
F:G:S    1   8.1233 193.81 2.58568  0.10783
G:A:S    1   5.5914 191.27 0.05372  0.81672
F:R:A:S  1   8.5419 194.22 3.00426  0.08305 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:A + F:G:A + R:G:A + F:R:S + F:G:S +
    F:A:S + R:A:S + F:R:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       5.5914 191.27
F:G:A    1   6.6924 190.37 1.10102  0.29404
R:G:A    1   5.6362 189.32 0.04486  0.83225
F:G:S    1   8.1405 191.82 2.54919  0.11035
F:R:A:S  1   8.5954 192.28 3.00405  0.08306 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - A:G:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:A + R:G:A + F:R:S + F:G:S + F:A:S +
    R:A:S + F:R:A:S
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       6.6924 190.37
R:G:A    1   6.9552 188.64 0.2628  0.60820
F:G:S    1   9.0266 190.71 2.3342  0.12656
F:R:A:S  1   9.7461 191.43 3.0537  0.08055 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G:F)
> drop1(fm, test="Chisq")
```

```
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:A + R:G:A + F:R:S + F:A:S + R:A:S +
    F:R:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       9.0266 190.71
F:G      1   9.5654 189.25 0.53889  0.46289
G:S      1  10.6144 190.30 1.58786  0.20763
R:G:A    1   9.1911 188.87 0.16455  0.68500
F:R:A:S  1  11.9996 191.68 2.97309  0.08466 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - G:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + G:S + A:S + F:R:A + R:G:A + F:R:S + F:A:S + R:A:S +
    F:R:A:S
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       9.5654 189.25
G:S      1  11.2930 188.97 1.7275   0.1887
R:G:A    1   9.7241 187.41 0.1587   0.6904
F:R:A:S  1  12.5377 190.22 2.9722   0.0847 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + A:S + F:R:A + R:G:A + F:R:S + F:A:S + R:A:S + F:R:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       11.293 188.97
R:G:A    1   11.424 187.11 0.13117   0.7172
F:R:A:S  1   14.265 189.95 2.97224   0.0847 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + A:S + F:R:A + R:G:A + F:R:S + F:A:S + R:A:S
      Df Deviance    AIC    LRT Pr(>Chi)
<none>      14.265 189.95
F:R:A   1   14.301 187.98 0.0355  0.85053
```

```
R:G:A   1   14.396 188.08 0.1312  0.71722
F:R:S   1   18.197 191.88 3.9316  0.04739 *
F:A:S   1   14.268 187.95 0.0033  0.95417
R:A:S   1   15.409 189.09 1.1442  0.28476
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + A:S + F:R:A + R:G:A + F:R:S + R:A:S
       Df Deviance    AIC    LRT Pr(>Chi)
<none>      14.268 187.95
F:R:A   1   14.306 185.99 0.0371   0.8473
R:G:A   1   14.400 186.08 0.1312   0.7172
F:R:S   1   18.298 189.98 4.0297   0.0447 *
R:A:S   1   15.414 187.10 1.1456   0.2845
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - A:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + A:S + R:G:A + F:R:S + R:A:S
       Df Deviance    AIC    LRT Pr(>Chi)
<none>      14.306 185.99
F:A     1   15.703 185.38 1.3977   0.2371
R:G:A   1   14.437 184.12 0.1312   0.7172
F:R:S   1   18.302 187.98 3.9961   0.0456 *
R:A:S   1   15.434 185.12 1.1278   0.2882
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + F:A + R:A + G:A + F:S +
    R:S + A:S + R:G:A + F:R:S
       Df Deviance    AIC    LRT Pr(>Chi)
<none>      15.434 185.12
F:A     1   16.689 184.37 1.2551  0.26258
A:S     1   19.339 187.02 3.9053  0.04813 *
R:G:A   1   15.565 183.25 0.1312  0.71722
F:R:S   1   19.248 186.93 3.8146  0.05081 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

```
> fm<- update(fm, .~. - A:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + R:A + G:A + F:S + R:S +
    A:S + R:G:A + F:R:S
      Df Deviance    AIC    LRT Pr(>Chi)
<none>      16.689 184.37
A:S    1   20.938 186.62 4.2490  0.03927 *
R:G:A  1   16.820 182.50 0.1312  0.71722
F:R:S  1   20.493 186.17 3.8041  0.05113 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + R:A + G:A + F:S + R:S +
    A:S + R:G:A
      Df Deviance    AIC    LRT  Pr(>Chi)
<none>      20.493 186.17
F:R    1   97.428 261.11 76.936 < 2.2e-16 ***
F:S    1   27.402 191.08  6.909  0.008575 **
R:S    1   21.063 184.74  0.570  0.450285
A:S    1   24.742 188.42  4.249  0.039274 *
R:G:A  1   20.624 184.31  0.131  0.717217
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + R:G + R:A + G:A + F:S + A:S +
    R:G:A
      Df Deviance    AIC    LRT Pr(>Chi)
<none>      21.063 184.74
F:R    1   97.432 259.11 76.370  < 2e-16 ***
F:S    1   27.406 189.09  6.343  0.01178 *
A:S    1   25.177 186.86  4.115  0.04251 *
R:G:A  1   21.194 182.88  0.131  0.71722
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> #Count ~ S + A + G + R + F + S:A + A:G + A:R + G:R + S:F + R:F + A:G:R
>
> pchisq(fm$deviance, fm$df.residual, lower.tail=FALSE)
[1] 0.333352
>
> fmtab <- xtabs(fitted(fm) ~F+R+G+A+S, dat)
> oddsratio_fn<- function(tab) {tab[1,1]*tab[2,2] / (tab[1,2]*tab[2,1])}
```

```
>
> apply(fmtab, c("G","A","R"), oddsratio_fn) #conditional odds SF
, , R = Black

       A
G             1       2
  Male   1.736655 1.736655
  Female 1.736655 1.736655

, , R = White

       A
G             1       2
  Male   1.736655 1.736655
  Female 1.736655 1.736655

> cfmtab2 <- apply(fmtab, c("F","S"),sum) #collapse over AGR
# == apply(xtabs(Count~F+R+G+A+S,dat), c("F","S"),sum)
> oddsratio_fn(cfmtab2) #marginal odds SF #not the same
[1] 1.746789
>
> cfmtab3 <- apply(fmtab, c("F","S","A","R"),sum) #collapse over G
> apply(cfmtab3, c("A","R"), oddsratio_fn) #SF
   R
A      Black    White
  1 1.736655 1.736655
  2 1.736655 1.736655
```

```
#################### (b) ########################

> fm <- glm(Count ~ F*R*G*A*S, dat, family = poisson)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F * R * G * A * S
        Df Deviance    AIC    LRT Pr(>Chi)
<none>       0.0000 201.68
F:R:G:A:S  1   3.6389 203.32 3.6389  0.05644 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:G:S +
    F:R:A:S + F:G:A:S + R:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       3.6389 203.32
F:R:G:A  1   3.7082 201.39 0.06931   0.7923
F:R:G:S  1   3.8445 201.53 0.20561   0.6502
F:R:A:S  1   6.2708 203.95 2.63190   0.1047
F:G:A:S  1   4.2643 201.95 0.62544   0.4290
R:G:A:S  1   3.9912 201.67 0.35232   0.5528
>
> #fm<- update(fm, .~. - A:G:R:F)
> #drop1(fm, test="Chisq")
>
> fm<- update(fm, .~. - S:G:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:A:S +
    F:G:A:S + R:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       3.8445 201.53
F:R:G:A  1   3.9883 199.67 0.14382  0.70452
F:R:A:S  1   6.6670 202.35 2.82253  0.09295 .
F:G:A:S  1   4.3396 200.02 0.49510  0.48166
R:G:A:S  1   4.0567 199.74 0.21222  0.64504
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:R)
> drop1(fm, test="Chisq")
Single term deletions
```

```
Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + R:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:A:S +
    F:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       4.0567 199.74
R:G:S    1   4.4423 198.12 0.38561  0.53462
F:R:G:A  1   4.1993 197.88 0.14256  0.70575
F:R:A:S  1   7.1681 200.85 3.11137  0.07775 .
F:G:A:S  1   4.6655 198.35 0.60880  0.43524
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:A:S + F:G:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       4.4423 198.12
F:R:G:A  1   4.5383 196.22 0.09598  0.75671
F:R:A:S  1   7.5507 199.23 3.10834  0.07789 .
F:G:A:S  1   5.2614 196.94 0.81903  0.36547
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:G:S + F:A:S + R:A:S + G:A:S + F:R:G:A + F:R:A:S
       Df Deviance    AIC     LRT Pr(>Chi)
<none>       5.2614 196.94
F:G:S    1   7.9923 197.67 2.73091  0.09842 .
G:A:S    1   5.3272 195.01 0.06583  0.79750
F:R:G:A  1   5.3687 195.05 0.10732  0.74321
F:R:A:S  1   8.3433 198.03 3.08198  0.07916 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:G)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
```

```
   F:G:S + F:A:S + R:A:S + F:R:G:A + F:R:A:S
      Df Deviance    AIC    LRT Pr(>Chi)
<none>       5.3272 195.01
F:G:S    1   8.0140 195.69 2.6868  0.10118
F:R:G:A  1   5.4302 193.11 0.1030  0.74826
F:R:A:S  1   8.4017 196.08 3.0745  0.07953 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + G:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
   F:A:S + R:A:S + F:R:G:A + F:R:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       8.0140 195.69
G:S      1   9.6608 195.34 1.64678   0.1994
F:R:G:A  1   8.0614 193.74 0.04736   0.8277
F:R:A:S  1  11.0094 196.69 2.99537   0.0835 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:G)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
   F:A:S + R:A:S + F:R:G:A + F:R:A:S
      Df Deviance    AIC     LRT Pr(>Chi)
<none>       9.6608 195.34
F:R:G:A  1   9.6907 193.37 0.02993   0.8627
F:R:A:S  1  12.6330 196.31 2.97224   0.0847 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
   F:S + R:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
   F:A:S + R:A:S + F:R:G:A
      Df Deviance    AIC    LRT Pr(>Chi)
<none>       12.633 196.31
F:R:S    1   16.565 198.25 3.9316  0.04739 *
F:A:S    1   12.636 194.32 0.0033  0.95417
R:A:S    1   13.777 195.46 1.1442  0.28476
F:R:G:A  1   12.663 194.34 0.0299  0.86266
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    R:A:S + F:R:G:A
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       12.636 194.32
F:R:S    1   16.666 196.35 4.0297   0.0447 *
R:A:S    1   13.782 193.46 1.1456   0.2845
F:R:G:A  1   12.666 192.35 0.0299   0.8627
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:S +
    F:R:G:A
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       13.782 193.46
A:S      1   17.706 195.39 3.9243  0.04759 *
F:R:S    1   17.615 195.30 3.8332  0.05025 .
F:R:G:A  1   13.812 191.49 0.0299  0.86266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:R:F)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + R:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:G:A
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       17.615 195.30
F:S      1   24.171 199.85 6.5556  0.01046 *
R:S      1   18.157 193.84 0.5414  0.46187
A:S      1   21.510 197.19 3.8952  0.04842 *
F:R:G:A  1   17.645 193.33 0.0299  0.86266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:R)
> drop1(fm, test="Chisq")
Single term deletions

Model:
```

```
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + A:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:G:A
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       18.157 193.84
F:S      1   24.174 197.86 6.0180  0.01416 *
A:S      1   21.946 195.63 3.7897  0.05157 .
F:R:G:A  1   18.186 191.87 0.0299  0.86266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> fm<- update(fm, .~. - S:A)
> drop1(fm, test="Chisq")
Single term deletions

Model:
Count ~ F + R + G + A + S + F:R + F:G + R:G + F:A + R:A + G:A +
    F:S + F:R:G + F:R:A + F:G:A + R:G:A + F:R:G:A
       Df Deviance    AIC    LRT Pr(>Chi)
<none>       21.946 195.63
F:S      1   28.421 200.10 6.4748  0.01094 *
F:R:G:A  1   21.976 193.66 0.0299  0.86266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> pchisq(fm$deviance, fm$df.residual, lower.tail=FALSE)
[1] 0.07972768
>
>
> fmtab <- xtabs(fitted(fm) ~F+R+G+A+S, dat)
> oddsratio_fn<- function(tab) {tab[1,1]*tab[2,2] / (tab[1,2]*tab[2,1])}
>
> apply(fmtab, c("G","A","R"), oddsratio_fn) #conditional odds SF
, , R = Black

       A
G             1       2
  Male   1.746789 1.746789
  Female 1.746789 1.746789

, , R = White

       A
G             1       2
  Male   1.746789 1.746789
  Female 1.746789 1.746789

>
> apply(fmtab, c("F","S"), sum) #collapse over GAR
             S
F             None Some
  BothParents 1120  109
  MotherOnly   200   34
> oddsratio_fn(apply(fmtab, c("F","S"), sum)) #collapse over GAR
[1] 1.746789
>
```

```
>
>
> (dat.logit <- cbind(expand.grid(A=A,G=G,R=R,F=F),
+                     SN = dat$Count[dat$S=="Some"], SY = dat$Count[dat$S=="None"]))
   A      G     R           F SN  SY
1  1   Male Black BothParents  2  27
2  2   Male Black BothParents  2  12
3  1 Female Black BothParents  4  23
4  2 Female Black BothParents  1   7
5  1   Male White BothParents 32 394
6  2   Male White BothParents 19 142
7  1 Female White BothParents 38 421
8  2 Female White BothParents 11  94
9  1   Male Black  MotherOnly  1  18
10 2   Male Black  MotherOnly  1  13
11 1 Female Black  MotherOnly  0  24
12 2 Female Black  MotherOnly  3   4
13 1   Male White  MotherOnly  6  48
14 2   Male White  MotherOnly  4  25
15 1 Female White  MotherOnly 15  55
16 2 Female White  MotherOnly  4  13
>
> fm.logit <- glm(cbind(SN, SY) ~ F, dat.logit, family =binomial)
> summary(fm.logit)

Call:
glm(formula = cbind(SN, SY) ~ F, family = binomial, data = dat.logit)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7452  -0.7685   0.1144   0.9863   1.8081

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.3297     0.1003 -23.220  < 2e-16 ***
FMotherOnly   0.5578     0.2109   2.645  0.00817 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 28.421  on 15  degrees of freedom
Residual deviance: 21.946  on 14  degrees of freedom
AIC: 75.655

Number of Fisher Scoring iterations: 4
```