Stat 8053, Fall 2013: Multinomial Logistic Regression (Faraway, Chap. 5)

Women's Labor Force Participation in Canada

```
library(car)
str(Womenlf)
```

```
'data.frame':        263 obs. of  4 variables:
 $ partic  : Factor w/ 3 levels "fulltime","not.work",..: 2 2 2 2 2 2 2 1 2 2 ...
 $ hincome : int  15 13 45 23 19 7 15 7 15 23 ...
 $ children: Factor w/ 2 levels "absent","present": 2 2 2 2 2 2 2 2 2 2 ...
 $ region  : Factor w/ 5 levels "Atlantic","BC",..: 3 3 3 3 3 3 3 3 3 3 ...
```

```
# reorder factor levels
Womenlf$partic <- factor(Womenlf$partic, levels=levels(Womenlf$partic)[c(2, 3, 1)])
Womenlf$region <- factor(Womenlf$region, levels=levels(Womenlf$region)[c(1, 4, 3,  5, 2)])
Womenlf$inc <- factor(ifelse(Womenlf$hincome < 14, "low", "high"), levels=c("low", "high"))
(tab <-ftable(xtabs(~ children + inc + partic, Womenlf)))
```

```
             partic not.work parttime fulltime
children inc
absent   low               13        3       27
         high              13        4       19
present  low               54       18       13
         high              75       17        7
```

```
round(100 * prop.table(tab, 1), 1)
```

```
             partic not.work parttime fulltime
children inc
absent   low             30.2      7.0     62.8
         high            36.1     11.1     52.8
present  low             63.5     21.2     15.3
         high            75.8     17.2      7.1
```

```
library(nnet) # multinom is in the nnet package
m1 <- multinom(partic ~ log(hincome) + children + region, data=Womenlf)
```

```
# weights:  24 (14 variable)
initial  value 288.935032
```

```
iter  10 value 209.772667
iter  20 value 208.795522
final  value 208.795515
converged
```

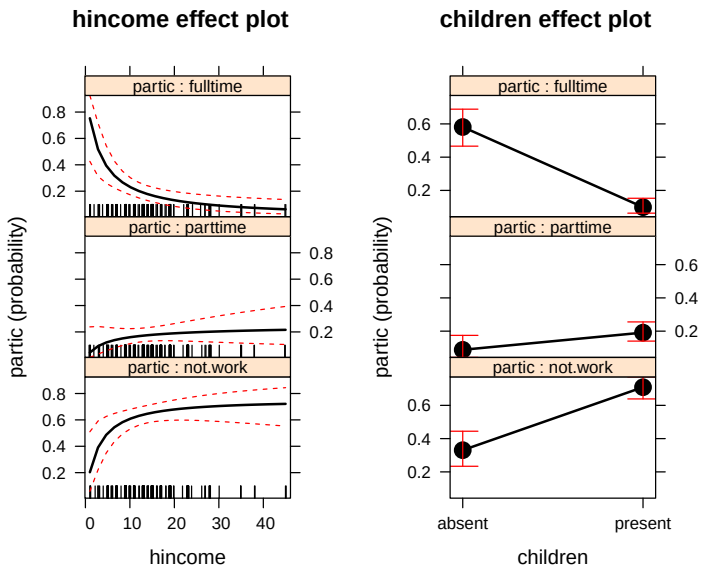
```
Anova(m1)
```

```
Analysis of Deviance Table (Type II tests)

Response: partic
             LR Chisq Df Pr(>Chisq)
log(hincome)     12.5  2     0.0019
children         63.2  2    1.9e-14
region            7.0  8     0.5320
```

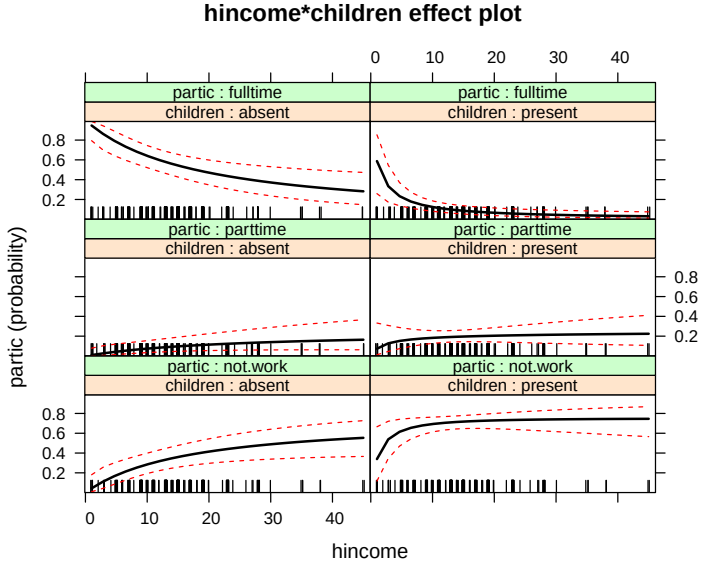
```
m2 <- update(m1, ~ . - region, trace=FALSE) # trace=FALSE shortens output
summary(m2)
```

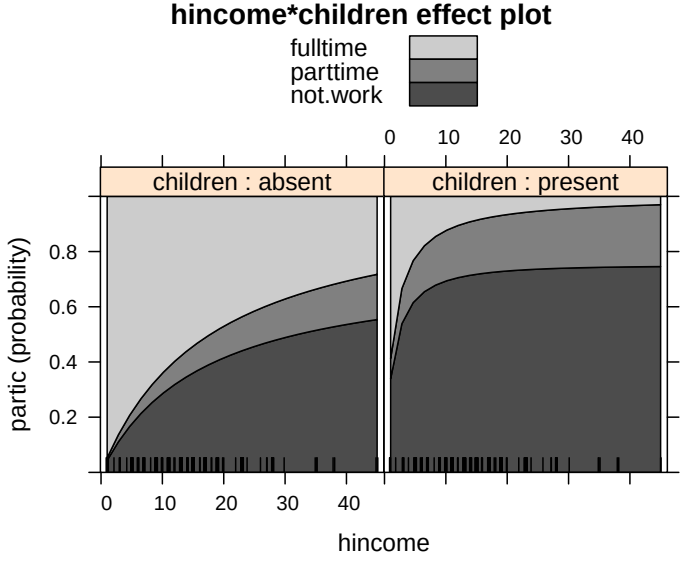
```
Call:
multinom(formula = partic ~ log(hincome) + children, data = Womenlf,
    trace = FALSE)

Coefficients:
         (Intercept) log(hincome) childrenpresent
parttime      -1.547      0.08667          0.0148
fulltime       3.080     -0.98585         -2.5307

Std. Errors:
         (Intercept) log(hincome) childrenpresent
parttime      1.0386       0.3492          0.4674
fulltime      0.8106       0.2960          0.3610

Residual Deviance: 424.6
AIC: 436.6
```

```
library(effects)
plot(allEffects(m2, xlevels=list(hincome=25)))
```




```
plot(effect("log(hincome)*children", m2, xlevels=list(hincome=25)))
plot(effect("log(hincome)*children", m2, xlevels=list(hincome=25)), style="stacked", colors=grey(c(.3, .5, .8)))
```





The `predict` method by default returns the category with the highest predicted probability:

```
(tab1 <- xtabs(~ partic + predict(m1), Womenlf))
```

```
          predict(m1)
partic     not.work parttime fulltime
  not.work      136        0       19
  parttime       36        0        6
  fulltime       23        0       43
```

```
accuracy <- function(tab) round( 100 * sum(diag(tab))/sum(tab), 1)
accuracy(tab1)
```

```
[1] 68.1
```

```
set.seed(1144)
dim(Womenlf)
```

```
[1] 263   5
```

```
construct <- sample(1:dim(Womenlf)[1], 150)
mconstruct <- update(m1, subset=construct, trace=FALSE)
accuracy( xtabs( ~ partic + predict(mconstruct), Womenlf[construct,]))
```

```
[1] 72
```

```
accuracy( xtabs( ~ partic + predict(mconstruct, Womenlf[-construct,]), Womenlf[-construct,]))
```

```
[1] 58.4
```

Compare to:

```
prop.table(xtabs(~ partic, Womenlf))
```

```
partic
not.work parttime fulltime
  0.5894   0.1597   0.2510
```

The `predict` with argument `type="probs"` returns vectors of estimated probabilities:

```
predict(m1, Womenlf[1:7,], type="probs")
```

```
  not.work parttime fulltime
1   0.7149   0.1939  0.09114
2   0.7054   0.1904  0.10416
3   0.7560   0.2128  0.03119
4   0.7367   0.2027  0.06055
5   0.7281   0.1991  0.07282
6   0.6480   0.1713  0.18069
7   0.7149   0.1939  0.09114
```

Similarly, the residuals return an $n \times m$ matrix whose elements are $y_{ij} - \hat{p}_{ij}$.

Central Nervous System Birth Defect Prevalence in South Wales

```
library(faraway)
data(cns)   # from faraway p. 103
cns
```

```
            Area NoCNS An Sp Other Water      Work
1        Cardiff  4091  5  9     5   110 NonManual
2        Newport  1515  1  7     0   100 NonManual
3        Swansea  2394  9  5     0    95 NonManual
4     GlamorganE  3163  9 14     3    42 NonManual
5     GlamorganW  1979  5 10     1    39 NonManual
6     GlamorganC  4838 11 12     2   161 NonManual
7      MonmouthV  2362  6  8     4    83 NonManual
8  MonmouthOther  1604  3  6     0   122 NonManual
9        Cardiff  9424 31 33    14   110    Manual
10       Newport  4610  3 15     6   100    Manual
11       Swansea  5526 19 30     4    95    Manual
12    GlamorganE 13217 55 71    19    42    Manual
13    GlamorganW  8195 30 44    10    39    Manual
14    GlamorganC  7803 25 28    12   161    Manual
15     MonmouthV  9962 36 37    13    83    Manual
16 MonmouthOther  3172  8 13     3   122    Manual
```

Predictors `Water` quality, and type of parent's `Work`. The response is given in 4 separate columns. This is grouped data.

```
prop.table(colSums(cns[, 2:5]))
```

```
   NoCNS       An       Sp    Other
0.991792 0.003028 0.004045 0.001135
```

Almost all cases are non-CNS. This suggests fitting two models, a logistic model for NoCNS vs. CNS, and then a multinomial logistic model for (An, Sp, Other)|CNS. For the logistic:

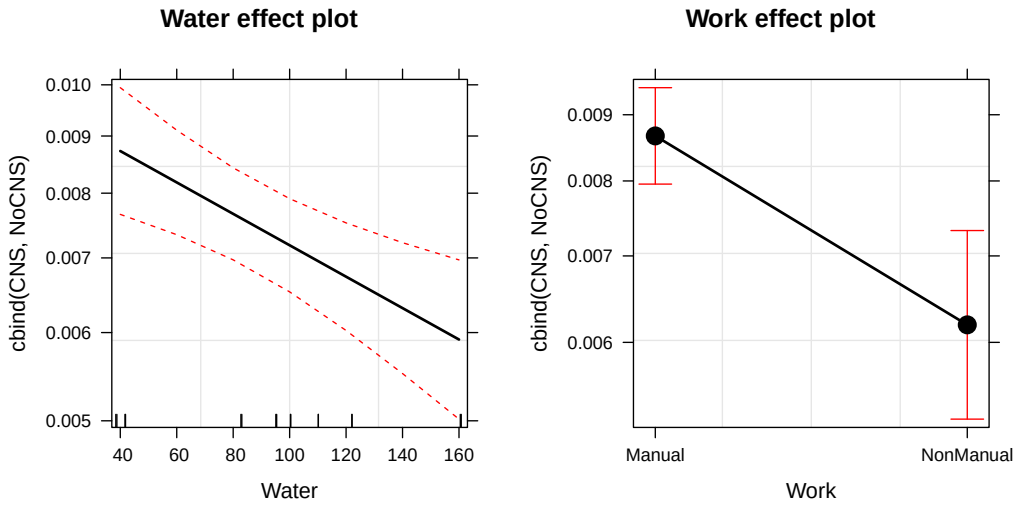
```
cns$CNS <- rowSums(cns[, 3:5])
fit1 <- glm(cbind(CNS, NoCNS) ~ Water + Work, binomial, cns)
summary(fit1)
```

```
Call:
glm(formula = cbind(CNS, NoCNS) ~ Water + Work, family = binomial,
    data = cns)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6557  -0.3018  -0.0313   0.5721   1.3300

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)   -4.432580   0.089789  -49.37  < 2e-16
Water         -0.003264   0.000968   -3.37  0.00075
WorkNonManual -0.339058   0.097094   -3.49  0.00048

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 41.047  on 15  degrees of freedom
Residual deviance: 12.363  on 13  degrees of freedom
AIC: 102.5

Number of Fisher Scoring iterations: 4
```

```
exp(confint(fit1)[-1,])
```

```
               2.5 % 97.5 %
Water         0.9948 0.9986
WorkNonManual 0.5868 0.8589
```

```
plot(allEffects(fit1), grid=TRUE)
```



The multinomial can be fit in two ways. Using a multivariate response

```
fit2 <- multinom(cbind(An, Sp, Other) ~ Water + Work, cns, trace=FALSE)
```

Or, using a univariate response with weights.

```
cns1 <- reshape(cns, varying= c("An", "Sp", "Other"), direction="long",
     v.name="y", timevar="outcome")
cns1$outcome <- factor(cns1$outcome, labels=c("An", "Sp", "Other"))
head(cns1)
```

```
          Area NoCNS Water      Work CNS outcome  y id
1.1    Cardiff  4091   110 NonManual  19      An  5  1
2.1    Newport  1515   100 NonManual   8      An  1  2
3.1    Swansea  2394    95 NonManual  14      An  9  3
4.1 GlamorganE  3163    42 NonManual  26      An  9  4
5.1 GlamorganW  1979    39 NonManual  16      An  5  5
6.1 GlamorganC  4838   161 NonManual  25      An 11  6
```

```
fit3 <- multinom(outcome ~ Water + Work, cns1, weights=y, trace=FALSE)
compareCoefs(fit2, fit3)
```

```
Call:
1:"multinom(formula = cbind(An, Sp, Other) ~ Water + Work, data = cns, trace = FALSE)"
2:"multinom(formula = outcome ~ Water + Work, data = cns1, weights = y, trace = FALSE)"
```

```
                      Est. 1     SE 1   Est. 2     SE 2
Sp:(Intercept)       0.37520  0.19003  0.37520  0.19003
Sp:Water            -0.00130  0.00203 -0.00130  0.00203
Sp:WorkNonManual     0.11576  0.20869  0.11576  0.20869
Other:(Intercept)   -1.12255  0.27956 -1.12255  0.27956
Other:Water          0.00218  0.00290  0.00218  0.00290
Other:WorkNonManual -0.27028  0.32472 -0.27028  0.32472
```

```
fit4 <- update(fit2, ~ 1)
anova(fit2, fit4)
```

```
Likelihood ratio tests of Multinomial Models

Response: cbind(An, Sp, Other)
         Model Resid. df Resid. Dev   Test    Df LR stat. Pr(Chi)
1            1        30       1374
2 Water + Work        26       1372 1 vs 2     4     2.93  0.5695
```

It isn't particularly surprising that the type of CNS cannot be predicted by `Water` and `Work`.