Multinomial logistic regression

- Below we use the multinom function from the nnet package to estimate a multinomial logistic regression model.
- There are other functions in other R packages capable of multinomial regression.
- We chose the multinom function because it does not require the data to be reshaped (as the mlogit package does) and to mirror the example code found in Hilbe's Logistic Regression Models.

Multinomial logistic regression

- Before running our model. We then choose the level of our outcome that we wish to use as our baseline and specify this in the relevel function. Then, we run our model using multinom.
- The multinom package does not include p-value calculation for the regression coefficients, so we calculate p-values using Wald tests (here z-tests).

```
ml$prog2 <- relevel(ml$prog, ref = "academic")
test <- multinom(prog2 ~ ses + write, data = ml)

 # weights:  15 (8 variable)
 initial  value 219.722458
 iter  10 value 179.982880
 final  value 179.981726
 converged
```

Multinomial Regression with R

```
summary(test)

 Call:
 multinom(formula = prog2 ~ ses + write, dat

 Coefficients:
          (Intercept) sesmiddle seshigh    w
 general        2.852   -0.5333 -1.1628 -0.0
 vocation       5.218    0.2914 -0.9827 -0.1
```

Multinomial Regression with R

```
....
 Std. Errors:
             (Intercept) sesmiddle seshigh   wr
 general           1.166    0.4437  0.5142 0.02
 vocation          1.164    0.4764  0.5956 0.02

 Residual Deviance: 360
 AIC: 376
```

Multinomial Regression with R

```
z <- summary(test)$coefficients/summary(test
z

          (Intercept) sesmiddle seshigh  wri
 general        2.445   -1.2018  -2.261 -2.7
 vocation       4.485    0.6117  -1.650 -5.1
```

Multinomial Regression with R

```
# 2-tailed z test
p <- (1 - pnorm(abs(z), 0, 1)) * 2
p

            (Intercept) sesmiddle seshigh
 general     1.448e-02    0.2294 0.02374 6.81
 vocation    7.299e-06    0.5408 0.09895 3.17
```

Multinomial Regression with R

- We first see that some output is generated by running the model, even though we are assigning the model to a new R object. This model-running output includes some iteration history and includes the final negative log-likelihood 179.981726.
- This value multiplied by two is then seen in the model summary as the Residual Deviance and it can be used in comparisons of nested models, but we won't show an example of comparing models on this page.

Multinomial Regression with R

- The model summary output has a block of coefficients and a block of standard errors. Each of these blocks has one row of values corresponding to a model equation.
- Focusing on the block of coefficients, we can look at the first row comparing prog = "general" to our baseline prog = "academic" and the second row comparing prog = "vocation" to our baseline prog = "academic".

Multinomial Regression with R

- If we consider our coefficients from the first row to be b_1 and our coefficients from the second row to be b_2, we can write our model equations:

$$ln\left(\frac{P(prog = general)}{P(prog = academic)}\right) = b_{10} + b_{11}(ses = 2) + b_{12}($$

$$ln\left(\frac{P(prog = vocation)}{P(prog = academic)}\right) = b_{20} + b_{21}(ses = 2) + b_{22}($$

Multinomial Regression with R

- A one-unit increase in the variable write is associated with the decrease in the log odds of being in general program vs. academic program in the amount of .058 (b_13).
- A one-unit increase in the variable write is associated with the decrease in the log odds of being in vocation program vs. academic program. in the amount of .1136 (b_23).

Multinomial Regression with R

- The log odds of being in general program vs. in academic program will decrease by 1.163 if moving from ses="low" to ses="high"(b_12).
- The log odds of being in general program vs. in academic program will decrease by 0.533 if moving from ses="low" to ses="middle"(b_11), although this coefficient is not significant.
- The log odds of being in vocation program vs. in academic program will decrease by 0.983 if moving from ses="low" to ses="high"(b_22).
- The log odds of being in vocation program vs. in academic program will increase by 0.291 if moving from ses="low" to ses="middle"(b_21).