Ordinal Dependent Variables

Week 18

Winnie Xia

What are ordinal dependent variables?

Ordered logit and probit regression

Ordered logit and probit in R

```
library(foreign)
dat <- read.dta("https://stats.idre.ucla.edu/stat/data/ologit.dta")
head(dat)
```

```
##             apply pared public  gpa
## 1     very likely     0      0 3.26
## 2 somewhat likely     1      0 3.21
## 3        unlikely     1      1 3.94
## 4 somewhat likely     0      0 2.81
## 5 somewhat likely     0      0 2.53
## 6        unlikely     0      1 2.59
```

Ordered logit and probit in R

Below we use the `polr` function from the MASS package to estimate an ordered logistic regression model.

```
library(MASS)
library(texreg)
m1 <- polr(apply ~ pared + public + gpa, data = dat,
           method = "logistic", Hess=TRUE)

m2 <- polr(apply ~ pared + public + gpa, data = dat,
           method = "probit", Hess=TRUE)

m3 <- glm(as.numeric(apply) ~ pared + public + gpa,
          data = dat)
```

Ordered logit and probit in R

```
##
## ============================================================
##                               Model 1      Model 2      Model 3
## ------------------------------------------------------------
## pared                          1.05 ***     0.60 ***     0.36 ***
##                               (0.27)       (0.16)       (0.09)
## public                        -0.06         0.01         0.02
##                               (0.30)       (0.17)       (0.10)
## gpa                            0.62 *       0.36 *       0.19 *
##                               (0.26)       (0.16)       (0.09)
## unlikely|somewhat likely       2.20 **      1.30 **
##                               (0.78)       (0.47)
## somewhat likely|very likely    4.30 ***     2.50 ***
##                               (0.80)       (0.48)
## (Intercept)                                               0.91 ***
##                                                          (0.25)
## ------------------------------------------------------------
## AIC                          727.02       727.50       798.34
## BIC                          746.98       747.45       818.29
## Log Likelihood              -358.51      -358.75      -394.17
## Deviance                     717.02       717.50       168.08
## Num. obs.                    400          400          400
## ============================================================
```

Interpretation

```
summary(m1)
```

```
## Call:
## polr(formula = apply ~ pared + public + gpa, data = dat, Hess = TRUE,
##     method = "logistic")
##
## Coefficients:
##           Value Std. Error t value
## pared   1.04769     0.2658  3.9418
## public -0.05879     0.2979 -0.1974
## gpa     0.61594     0.2606  2.3632
##
## Intercepts:
##                             Value   Std. Error t value
## unlikely|somewhat likely     2.2039  0.7795     2.8272
## somewhat likely|very likely  4.2994  0.8043     5.3453
##
## Residual Deviance: 717.0249
## AIC: 727.0249
```

Interpretation

We often interpret models by employing prediction based on scenarios.

The Parallel Regression Assumption

- Assessing the parallel regression assumption