

Limited Dependent Variables

Week 20

Winnie Xia

Limited Dependent Variable

- What is a limited dependent variable? **Limited dependent variables** are dependent variables that have limited ranges, usually either **discontinuous** or **range bounded**.

Censoring and Truncation

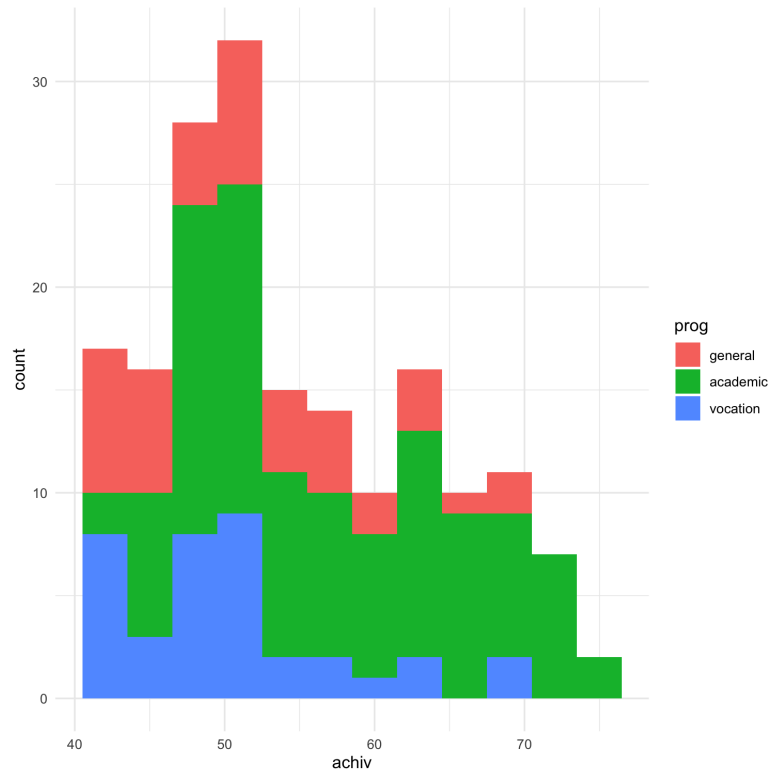
- Censoring is when the limit observations are in the sample.
- Truncation is when the observations are not in the sample.

Truncated Regression

- Truncated regression addresses the bias introduced when using OLS regression with truncated data.

Truncated Regression in R

```
library(truncreg)
dat <- read.dta("https://stats.idre.ucla.edu/stat/data/truncreg.dta")
ggplot(dat, aes(achiv, fill = prog)) +
  geom_histogram(binwidth=3) +
  theme_minimal()
```



Truncated Regression in R

```
m <- truncreg(achiv ~ langscore + prog, data = dat,  
              point = 40, direction = "left")  
screenreg(m)
```

```
##  
## =====  
##                               Model 1  
## -----  
## (Intercept)           11.30  
##                      (6.77)  
## langscore              0.71 ***  
##                      (0.11)  
## progacademic           4.06 *  
##                      (2.05)  
## progvocation          -1.14  
##                      (2.67)  
## sigma                 8.75 ***  
##                      (0.67)  
## -----  
## Num. obs.             178  
## Log Likelihood        -591.31  
## AIC                   1192.62  
## BIC                   1208.53  
## =====
```

Tobit Model

The tobit model, also called a censored regression model, is designed to estimate linear relationships between variables when there is either left- or right-censoring in the dependent variable (also known as censoring from below and above, respectively).

Tobit Model in R

```
library(VGAM)
acad_apt = read_csv("https://stats.idre.ucla.edu/stat/data/tobit.csv") %
  mutate(prog = factor(prog, labels = c('acad',
                                         'general', 'vocational')))
```


Tobit Model in R

```
m2 <- vglm/apt ~ read + math + prog, tobit(Upper = 800), data = acad_apl  
screenreg(m2)
```

```
##  
## =====  
##                      Model 1  
## -----  
## (Intercept):1      209.56 ***  
##                   (32.55)  
## (Intercept):2       4.18 ***  
##                   (0.05)  
## read                2.70 ***  
##                   (0.62)  
## math                5.91 ***  
##                   (0.71)  
## proggeneral        -12.71  
##                   (12.41)  
## progvocational     -46.14 ***  
##                   (13.71)  
## -----  
## Log Likelihood    -1041.06  
## DF                 394  
## Num. obs.         400  
## =====
```

Heckit Model

- Heckit model can be used when there is incidental truncation, i.e., there is variable that explains whether we have observations for an outcome variable.
- **Example:**
 - wage for those who have jobs, another variable that indicates whether someone has a job;
 - grades for those who go to universities, another variable that indicates whether someone goes to university or not.
- Two-Stage least squares, **a statistical technique that is used in the analysis of structural equations.**