EC508: Econometrics
Beyond Binary Outcome

Jean-Jacques Forneron

Spring, 2023

Boston University

Sometimes we face more than one option

Multiple Outcomes

1. Multinomial choice:
    - Y_i: white, orange, blue car
    - X_i: car characteristics

    Use multinomial logit and probit models
2. Ordered Response
    - Y_i: high-school degree (0), college education (1), graduate school (2)
    - X_i: student characteristics

    Use ordered probit model

Multinomial Logit model

- Applies to many settings:
  - Choosing a car amongst multiple colors/brands/...
  - Choosing a leisure activity: climbing/reading/opera/...
  - etc.
- Used in economics to model qualitative decisions but also in marketing to understand consumer behaviour
- Each choice j has a utility $U_i(j)$:

$$U_i(j) = X_j' \beta_j + u_{i,j}$$

  assume a Gumbel, aka extreme value, distribution: $F(u_{i,j}) = \exp(-\exp(-u_{i,j}))$
- Then

$$\mathbb{P}(j \text{ is chosen}|X) = \frac{\exp(X_j' \beta_j)}{\sum_{\ell=1}^{J} \exp(X_\ell' \beta_\ell)}$$

Multinomial Logit model, cont'd

- Likelihood function:

$$L(\beta|Y,X) = \prod_{i=1}^{n} \left(\prod_{j=1}^{J} \left[\frac{\exp(X_j'\beta_j)}{\sum_{\ell=1}^{J} \exp(X_\ell'\beta_\ell)} \right]^{\mathbb{1}_{Y_{i,j}=j}} \right)$$

- Log-likelihood:

$$\ell(\beta|Y,X) = \sum_{i=1}^{n} \left(\sum_{j=1}^{J} \mathbb{1}_{Y_{i,j}=j} \log \left[\frac{\exp(X_j'\beta_j)}{\sum_{\ell=1}^{J} \exp(X_\ell'\beta_\ell)} \right] \right)$$

- This function is very easy to maximize numerically

Sample selection model

- Suppose $Y_i^\star = X_i'\beta + u_i$ are e.g. wages
- We only observe:

$$D_i = \mathbb{1}\{X_i'\gamma + W_i'\delta + v_i > 0\}$$
$$Y_i = Y_i^\star \text{ if } D_i = 1, \quad Y_i = 0 \text{ otherwise}$$

- Suppose we only keep Y_i when $D_i = 1$ and regress it on X_i, then

$$\mathbb{E}(Y_i^\star | X_i, D_i = 1) = X_i'\beta + \mathbb{E}(u_i | X_i, D_i = 1)$$

- Now note that:

$$\mathbb{E}(u_i | X_i, D_i = 1) = \mathbb{E}(u_i | X_i, v_i > -X_i'\gamma - W_i'\delta)$$

- If v_i and u_i are correlated, then this expectation is $\neq 0$; the estimator is biased and inconsistent

Sample selection model, cont'd

- We can use the sample selection model to model labor participation problems:
    - D_i: decision to participate in the labour market
    - Y_i: observed wage if the individual enters the labour market
- If we want to study wages Y_i we have to account for the labour participation decision
- This problem is very common in economic settings: e.g. firms entering a market (country, state, city); consumers entering a program (club membership); etc.
- This model can be estimated by MLE if we make some assumptions about u_i, v_i

Sample selection model, cont'd

- Suppose

$$u_i, v_i \sim \mathcal{N}\left(0, \begin{pmatrix} \sigma_u^2 & \rho\sigma_u \\ \rho\sigma_u & 1 \end{pmatrix}\right)$$

  independent of X_i, W_i

- Under this assumption:

$$\mathbb{E}(u_i | X_i, W_i, D_i = 1) = \rho\sigma_u \lambda(-[X_i'\gamma + W_i'\delta])$$

- λ is the inverse Mill's ratio evaluated at the z-score for D_i:

$$\lambda(x) = \frac{\phi(x)}{1 - \Phi(x)}$$

- Overall we get:

$$\mathbb{E}(Y_i | X_i, D_i = 1) = X_i'\beta + \rho\sigma_u \lambda(-[X_i'\gamma + W_i'\delta])$$

- This is fairly easy to estimate with non-linear least-squares, regress: estimate γ, δ using a probit; then regress the Y_i for which $D_i = 1$ on X_i and $\lambda(-[X_i'\hat{\gamma} + W_i'\hat{\delta}])$ using OLS
- This two-step (probit+OLS) approach is very simple but MLE is efficient...
- The likelihood is complicated, it involves integrals
- In R, you can use the *sampleSelection* package to estimate the sample selection model
- Remark: for the general model we need W_i not in the main regression to identify β

Example: labour force participation and children

- Mroz (1987) analyzed female labour supply
- Labour force participation (lfp) is modelled by a quadratic polynomial in age, family income in 1975 dollars, presence of children and education in years.
- Wage equation: a quadratic polynomial in experience, education in years and residence in a big city

Application to the Boston Housing Data: Probit

```r
library(sampleSelection); data( "Mroz87" )
Mroz87$kids <- ( Mroz87$kids5 + Mroz87$kids618 > 0 )

TwoStep <- selection( lfp ~ age + I( age^2 ) + faminc + kids + educ,wage ~ exper + I( exper^2 ) + educ + city,data = Mroz87, method = "2step" )

MLE <- selection( lfp ~ age + I( age^2 ) + faminc + kids + educ,   wage ~ exper + I( exper^2 ) + educ + city, data = Mroz87, maxMethod = "BHHH", iterlim = 500 )

summary(TwoStep); summary(MLE)
```

Two-Step Estimates

	Estimate	Std. Error	t value	Pr(>—t—)
(Intercept)	−4.157	1.402	−2.965	0.003
age	0.185	0.066	2.810	0.005
I(age^2)	−0.002	0.001	−3.136	0.002
faminc	0.000	0.000	1.089	0.277
kids	−0.449	0.131	−3.430	0.001
educ	0.098	0.023	4.272	0.000
(Intercept)	−0.971	2.059	−0.472	0.637
exper	0.021	0.062	0.337	0.736
I(exper^2)	0.000	0.002	0.073	0.942
educ	0.417	0.100	4.160	0.000
city	0.444	0.316	1.405	0.160
invMillsRatio	−1.098	1.266	−0.867	0.386
sigma	3.200			
rho	−0.343			

Maximum Likelihood Estimates

	Estimate	Std. Error	t value	Pr(>—t—)
(Intercept)	−4.120	1.410	−2.921	0.004
age	0.184	0.066	2.795	0.005
I(age^2)	−0.002	0.001	−3.115	0.002
faminc	0.000	0.000	1.459	0.145
kids	−0.451	0.137	−3.298	0.001
educ	0.095	0.024	3.973	0.000
(Intercept)	−1.954	1.675	−1.167	0.244
exper	0.028	0.075	0.377	0.706
I(exper^2)	0.000	0.002	−0.049	0.961
educ	0.456	0.096	4.754	0.000
city	0.445	0.426	1.046	0.296
sigma	3.104	0.084	37.088	0.000
rho	−0.133	0.223	−0.598	0.550