Lab 4
Mode Choice (1)

Yufeng Zhang
zhan4879@umn.edu

CEGE-3201: Transportation Engineering

February 14, 2019

University of Minnesota

Announcements

- Transpo Talk: Feb 20(Wed.) 12:30 - 1:30 pm Room 780B
- My office: Room 175
- Report grading

Objectives

- Become familiar with transportation mode choice modeling
- Learn how to generate logit models in R
- Interpret regression results and construct utility functions
- Calculate probabilities for choosing different modes

Discrete Choice

Example: Travel mode to work

- Decision maker: an individual worker
- Choice set: ①Drive alone (DA), ②Carpool (CA), ③Transit (TR)
- Utility function: $U(x) = U(DA, CP, TR)$
- Outcomes: $\{1, 0, 0\}, \{0, 1, 0\}, \{0, 0, 1\}$

Decision maker's rule: **Maximize Utility**

- If $U(CP) > U(DA), U(CP) > U(TR)$
- Carpool
- If $U(TR) > U(CP), U(TR) > U(DA)$
- Transit

Discrete Choice

Example: Travel mode to work

- Decision maker: an individual worker
- Choice set: ①Drive alone (DA), ②Carpool (CA), ③Transit (TR)
- Utility function: $U(x) = U(DA, CP, TR)$
- Outcomes: $\{1, 0, 0\}, \{0, 1, 0\}, \{0, 0, 1\}$

Decision maker's rule: **Maximize Utility**

- If $U(CP) > U(DA), U(CP) > U(TR)$
- Carpool
- If $U(TR) > U(CP), U(TR) > U(DA)$
- Transit

Discrete Choice

Example: Travel mode to work

- Decision maker: an individual worker
- Choice set: ①Drive alone (DA), ②Carpool (CA), ③Transit (TR)
- Utility function: $U(x) = U(DA, CP, TR)$
- Outcomes: $\{1, 0, 0\}, \{0, 1, 0\}, \{0, 0, 1\}$

Decision maker's rule: **Maximize Utility**

- If $U(CP) > U(DA), U(CP) > U(TR)$
- Carpool
- If $U(TR) > U(CP), U(TR) > U(DA)$
- Transit

Discrete Choice

Example: Travel mode to work

- Decision maker: an individual worker
- Choice set: ①Drive alone (DA), ②Carpool (CA), ③Transit (TR)
- Utility function: $U(x) = U(DA, CP, TR)$
- Outcomes: $\{1, 0, 0\}, \{0, 1, 0\}, \{0, 0, 1\}$

Decision maker's rule: **Maximize Utility**

- If $U(CP) > U(DA), U(CP) > U(TR)$
- Carpool
- If $U(TR) > U(CP), U(TR) > U(DA)$
- Transit

Utility Function

What attributes would you consider when choosing travel mode?
$U(DA) = U$(travel time, parking cost, toll, ...)
$U(CP) = U$(travel time, convenience, privacy, ...)
$U(TR) = U$(travel time, transfer, walk time, ...)

Let's assume linear:

$$U_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ...$$

However, there are unobserved attributes/ measurement errors...

$$U_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ... + \epsilon_i$$

Multinomial Logit Model (MNL)

If we further assume that ϵ_i's are i.i.d and follow the Gumbel distribution, then the probability of choosing an alternative i from choice set C is:

$$P_i = \frac{e^{V_i}}{\sum_{j \in C} e^{V_j}}$$

- Traditional choice model
- Computes probability of one choice relative to another
- In transportation, used in mode choice and route choice modeling

Excercise

We have utility functions of walking and riding a scooter.
$V_{walk} = -2 - 0.2 \times T_{walk}$, and
$V_{scooter} = -1.5 - 0.1 \times T_{scooter} - 0.5 \times C_{rental}$.
T_{walk} and $T_{scooter}$ are travel times for walking and riding a scooter, respectively. C_{rental} is the monetary cost for using a scooter. Chris lives in the Centennial Hall. To come to the Civil Engineering Building, it takes 8 minutes by walking, or 3 minutes of riding a scooter plus a cost of 5 minutes equivalent of time.
What is the probability that Chris is going to ride a scooter?

Excercise

We have utility functions of walking and riding a scooter.
$V_{walk} = -2 - 0.2 \times T_{walk}$, and
$V_{scooter} = -1.5 - 0.1 \times T_{scooter} - 0.5 \times C_{rental}$.
T_{walk} and $T_{scooter}$ are travel times for walking and riding a scooter, respectively. C_{rental} is the monetary cost for using a scooter. Chris lives in the Centennial Hall. To come to the Civil Engineering Building, it takes 8 minutes by walking, or 3 minutes of riding a scooter plus a cost of 5 minutes equivalent of time.
What is the probability that Chris is going to ride a scooter?

Step 1: Calculating utilities

$V_{walk} = -2 - 0.2 \times 8 = -3.6$

$V_{scooter} = -1.5 - 0.1 \times 3 - 0.5 \times 5 = -4.3$

Step 2: Calculating probability

$$P_{scooter} = \frac{e^{V_{scooter}}}{e^{V_{scooter}} + e^{V_{walking}}} = \frac{e^{-4.3}}{e^{-4.3} + e^{-3.6}} = 0.33$$

Back to the Utility Function

In practice, a more interesting problem is to estimate parameters in the utility functions given observations of decision makers' choices. For each alternative i:

$$U_i = ASC_i + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \ldots$$

- ASC_i: Alternative specific constant (ASC)
- $\beta_1, \beta_2, \beta_3$: Coefficient to be estimated
- X_{i1}, X_{i2}, X_{i3}: Attributes associated with alternative i

Estimation

Maximum likelihood (ML) method.

$$L(\beta) = \prod_n \prod_j P(y_n = j|x_n)^{D_{nj}}$$

$$\text{Log(L}(\beta)) = \sum_n \sum_j D_{nj} LnP(y_n = j|x_n)$$

$$\text{Log(L}(\beta)) = \sum_n (\sum_j D_{nj}(x_{nj}^T \beta) - Ln(\sum_k \exp(x_{nk}^T \beta))$$

Nonlinear programming. N-R(Newton-Raphson), Gauss-Newton(G-N).
R is here to help you.

Estimating Parameters Using R

Data sets

- R package - "mlogit"
- Two travel mode data sets — "Mode" and "ModeCanada"

How many choice situations are there in data sets?
How many alternatives are there in data sets?
What attributes are recorded in data sets?

Construct Utility Functions

```
Coefficients :
                      Estimate  Std. Error  z-value  Pr(>|z|)
train:(intercept) -2.22533181  0.14607971 -15.2337 < 2.2e-16 ***
air:(intercept)   -1.65806769  0.49002931  -3.3836 0.0007154 ***
bus:(intercept)   -4.12693724  0.58645281  -7.0371 1.963e-12 ***
cost              -0.02721251  0.00431696  -6.3036 2.907e-10 ***
ivt                0.00138855  0.00096601   1.4374 0.1506023
train:dist         0.00319847  0.00051456   6.2159 5.103e-10 ***
air:dist           0.01117571  0.00060542  18.4595 < 2.2e-16 ***
bus:dist          -0.00532501  0.00256158  -2.0788 0.0376361 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -3070.3
McFadden R^2:  0.29663
Likelihood ratio test : chisq = 2589.6 (p.value = < 2.22e-16)
```

Figure: Multinomial Logit Regression Results

Construct Utility Functions (Cont.)

"Car" is used as a reference mode.
Read carefully about these utility functions.
Identify alternative specific variables, generic variables, and ASC's.

$$
\begin{aligned}
V_{train} &= -2.225 - 0.027X_{COST} + 0.001X_{IVT} + 0.003X_{DIST} \\
V_{air} &= -1.658 - 0.027X_{COST} + 0.001X_{IVT} + 0.011X_{DIST} \\
V_{bus} &= -4.127 - 0.027X_{COST} + 0.001X_{IVT} - 0.005X_{DIST} \\
V_{car} &= +0.000 - 0.027X_{COST} + 0.001X_{IVT}
\end{aligned}
$$

Calculate Probabilities of Choosing Each Mode

See exercise page.

Practice in R.

- Package "mlogit"
- Build regression models
- Document "Mode Choice (1).pdf" to help you