Mode Choice

EWCE 351
Transportation Engineering

Mode choice

- What portion & how many passenger trips between origin *i* & destination *j* are made by each of the following modes?
  - Motorized
    - Automobile
      - By occupancy: drive-alone, two-person, three+-person
      - By facility usage: general purpose, toll lanes, etc.
    - Public transit
      - By type: local bus, express bus, bus rapid transit, light rail, commuter rail
      - By access/egress mode: walk, drive
  - Non-motorized
    - Walk
    - Bicycle

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

What factors affect your choice of mode?

- Mode availability
- Trip purpose
- Level-of-service
  - Travel time: in-vehicle, wait/transfer, reliability
  - Trip cost: fuel, parking, tolls
  - Route attributes: pleasant scenery, turns, stops, terrain
- Traveler characteristics: income, etc.
- Area characteristics: density, etc.

How might you estimate mode choices?

- Heuristics
  - Based on trip distances & travel times
  - e.g., walking for all trips < 0.5 miles; driving for all trips > 3 miles
- Fixed percentages
  - Based on existing mode shares (from data collection)
  - e.g., 7% walk, 2% bicycle, 14% transit, 77% auto
- Combination:
  - Based on existing mode shares for trips of varying distance/duration
  - e.g., for trips < 0.5 miles: 75% walk, 10% bike, 0% transit, 15% auto

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

Mode choice model

- Probabilistic
- Non-linear regression on independent variables
  - Level-of-service, traveler characteristics, area characteristics
- Dependent variable is not mode choice but unobserved mode-specific variable:
  - "Utility"
- Assumptions based on error distributions & equation functional form:
  - Multinomial logit model
  - Nested logit model

} Discrete choice models

Discrete choice behavioral theory

- Foundations
  - Mathematical psychology
  - Microeconomics
- Four elements
  - Decision maker
  - Alternatives
  - Attributes of alternatives
  - Decision rule

Discrete choice behavioral theory

- Alternatives
  - Choice set: categorical, mutually exclusive, complete
  - Feasible and known
  - e.g., drive alone, carpool, bus, bicycle, walk

- Attributes of alternatives
  - Characteristics of the alternatives
  - e.g., travel time, cost, comfort, etc.

Discrete choice behavioral theory

- Decision maker
  - Rational actor: aware of alternatives and their attributes
  - Consistent and transitive preferences
    - Will make same choice under same situation
    - Alternatives can be ranked in order of preference
- Decision rule
  - How alternative is chosen from choice set of alternatives
  - Dominance
  - Satisfaction
  - **Utility maximization**

Utility maximization

- Utility (U)
  - Measure of satisfaction from consumption decision (i.e., consuming travel)
  - Latent variable: we cannot observe, we can only infer
  - No natural scale or fixed zero point: only differences matter

- Decision rule
  - Select the alternative that provides the highest (most positive) utility

Utility maximization

- Example

Alternative	Time	Cost
Car	5 minutes	$5.00
Bus	15 minutes	$2.00

- Utility function: $U_k = (-0.2) \times \text{time}_k + (-1) \times \text{cost}_k$
- Which alternative will be chosen?
  - $U_{car} = (-0.2) \times (5) + (-1) \times (5) = (-1) + (-5) = -6$
  - $U_{bus} = (-0.2) \times (15) + (-1) \times (2) = (-3) + (-2) = -5$
  - $(U_{bus} = -5) > (U_{car} = -6)$ → choose bus

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Utility maximization

- Example

Alternative	Time	Cost
Car	5 minutes	$5.00
Bus	15 minutes	$2.00

- Utility function: $U_k = (-0.5) \times \text{time}_k + (-1) \times \text{cost}_k$
- Which alternative will be chosen?
  - $U_{car} = (-0.5) \times (5) + (-1) \times (5) = (-2.5) + (-5) = -7.5$
  - $U_{bus} = (-0.5) \times (15) + (-1) \times (2) = (-7.5) + (-2) = -9.5$
  - $(U_{car} = -7.5) > (U_{bus} = -9.5)$ → choose car

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Utility maximization

- Really analyzing a three-alternative problem
- Example

Alternative	Time	Cost	Benefit
Car to shop	5 minutes	$5.00	Yes
Bus to shop	15 minutes	$2.00	Yes
Stay at home	0 minutes	$0.00	No

- Utility function: $U_k = (-0.5) \times \text{time}_k + (-1) \times \text{cost}_k + (20) \times \text{benefit}_k$
- Which alternative will be chosen?
  - $U_{car} = (-0.5) \times (5) + (-1) \times (5) + (20) \times (1) = (-2.5) + (-5) + (20) = 12.5$
  - $U_{bus} = (-0.5) \times (15) + (-1) \times (2) + (20) \times (1) = (-7.5) + (-2) + (20) = 10.5$
  - $U_{home} = (-0.5) \times (0) + (-1) \times (5) + (20) \times (0) = (0) + (0) + (0) = 0.0$
  - $U_{car} > U_{bus} > U_{home}$ → choose car

CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Random utility maximization (RUM)

- Analyst/modeler/planner cannot see inside travelers' heads
  - Decision-makers still assumed to choose highest utility alternative
- Must treat utility as an unobserved random variable
- Sources of randomness:
  - Unobserved attributes
    - We may not be able to measure all of the attributes considered
  - Measurement errors
    - Attributes measured are imperfect approximations of attributes considered
  - Unobserved taste variation
    - People have different preferences, make different tradeoffs between attributes

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

Random utility maximization

- Answer is probabilistic, not deterministic
- Split utility equation into measured and unmeasured elements $U_k = V_k + \varepsilon_k$
  - Deterministic/systematic/observable/measurable utility

$$V_k = \sum_m \beta_m \times X_{mk}$$

  - Stochastic/random/unobservable/unmeasurable utility

$$\varepsilon_k \qquad with\ expected\ mean = 0$$

- Mathematical form of the model depends on assumptions about the error term ε_k

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

Multinomial logit (MNL) model

$$P_n(k \mid C_n) = P_{kn} = \exp(V_k) \div \sum_{l \text{ in } Cn} \exp(V_l)$$

- For two alternatives
  - $P_{1n} = \exp(V_1) \div [\exp(V_1) + \exp(V_2)]$
  - $P_{2n} = \exp(V_2) \div [\exp(V_1) + \exp(V_2)]$

- For three alternatives
  - $P_{1n} = \exp(V_1) \div [\exp(V_1) + \exp(V_2) + \exp(V_3)]$
  - $P_{2n} = \exp(V_2) \div [\exp(V_1) + \exp(V_2) + \exp(V_3)]$
  - $P_{3n} = \exp(V_3) \div [\exp(V_1) + \exp(V_2) + \exp(V_3)]$

Why "logit"?

- Definitions
  - Probability p — the likelihood of choosing alternative k (0–1)
  - Odds = $p_1 \div (1 - p_1)$ — the ratio of probability(yes) to probability(no)
  - Log-odds = ln (odds) — the log of the odds (a.k.a., the "logit")

- $p = \exp(V_1) \div (\exp(V_1) + \exp(V_2)) = \exp(V_{diff}) \div (\exp(V_{diff}) + 1)$
- $$\begin{aligned} \text{Odds} &= [\exp(V_{diff}) \div (\exp(V_{diff}) + 1)] \div [1 - \{\exp(V_{diff}) \div (\exp(V_{diff}) + 1)\}] \\ &= [\exp(V_{diff}) \div (\exp(V_{diff}) + 1)] \div [1 \div (\exp(V_{diff}) + 1)] \\ &= \exp(V_{diff}) \div 1 = \exp(V_{diff}) \end{aligned}$$
- Log-odds = $\ln(\exp(V_{diff})) = V_{diff}$ = logit

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

MNL model example

Alternative	Time	Cost
Car	5 minutes	$5.00
Bus	15 minutes	$2.00

$$V_k = (-0.2) \times \text{time}_k + (-1) \times \text{cost}_k$$

- Probabilities of choosing each alternative
  - $V_{car} = (-0.2) \times (5) + (-1) \times (5) = (-1) + (-5) = -6$
  - $V_{bus} = (-0.2) \times (15) + (-1) \times (2) = (-3) + (-2) = -5$
  - $\exp(V_{car}) = e^{-6} = 0.00248$
  - $\exp(V_{bus}) = e^{-5} = 0.00674$
  - $\text{Prob(car)} = 0.00248 \div (0.00248 + 0.00674) = 0.00248 \div 0.00922 = 0.269$
  - $\text{Prob(bus)} = 0.00674 \div (0.00248 + 0.00674) = 0.00674 \div 0.00922 = 0.731$

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

MNL model example

Alternative	Time	Cost
Car	5 minutes	$5.00
Bus 1 (blue)	15 minutes	$2.00
Bus 2 (red)	15 minutes	$2.00

$$V_k = (-0.2) \times \text{time}_k + (-1) \times \text{cost}_k$$

- Probabilities of choosing each alternative
  - $V_{car} = -6$, $V_{bus1} = -5$, $V_{bus2} = -5$
  - $\exp(V_{car}) = 0.00248$, $\exp(V_{bus1}) = 0.00674$, $\exp(V_{bus2}) = 0.00674$,
    $\exp(V_{car}) + \exp(V_{bus1}) + \exp(V_{bus2}) = 0.00248 + 0.00674 + 0.00674 = 0.01595$
  - Prob(car) = $0.00248 \div 0.01595$ $= 0.155$
  - Prob(bus1) = Prob(bus2) = $0.00674 \div 0.01595$ $= 0.422$

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

MNL model example

Alternative	Time	Cost
Car 1 (Uber)	5 minutes	$5.00
Car 2 (Lyft)	5 minutes	$5.00
Bus	15 minutes	$2.00

$$V_k = (-0.2) \times \text{time}_k + (-1) \times \text{cost}_k$$

- Probabilities of choosing each alternative
  - $V_{car1} = -6,\ V_{car2} = -6,\ V_{bus} = -5$
  - $\exp(V_{car1}) = 0.00248,\ \exp(V_{car2}) = 0.00248,\ \exp(V_{bus}) = 0.00674,$
    $\exp(V_{car}) + \exp(V_{bus1}) + \exp(V_{bus2}) = 0.00248 + 0.00248 + 0.00674 = 0.01170$
  - $\text{Prob(car1)} = \text{Prob(car2)} = 0.00248 \div 0.01170 = 0.212$
  - $\text{Prob(bus)} = 0.00674 \div 0.01170 = 0.576$

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Independence from irrelevant alternatives (IIA)

- Example 1: car and bus alternatives
  - 27% car, 73% bus
- Example 2: car and two (nearly identical) bus alternatives
  - Red bus / blue bus paradox
  - 16% car, 84% bus (total)
- Example 3: bus and two (nearly identical) car alternatives
  - Let's call this the "Uber/Lyft" paradox
  - 42% car (total), 58% bus
- We might expect car and bus mode shares to remain ~ like in Example 1

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Independence from irrelevant alternatives (IIA)

- In general, for a given decision-maker n,
  the ratio of the choice probabilities of two alternatives P_{kn} / P_{ln}
  is unaffected by addition, subtraction, or changes
  to any other alternative m.

- Why?
  - Independently distributed errors ε_k
  - Unobserved utility for both bus/car alternatives is likely highly correlated
    (e.g., similar comfort levels, crowding, user preferences, etc.)

- How to avoid?
  - Nested logit model

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

Nested logit model

- Assign alternatives to nests, select (or estimate) a nesting parameter m
  - m is based on (1 – correlation of alternatives within nest)
  - m ranges from near 0 (completely similar, like one alternative) to near 1 (completely different, like red/blue buses or Uber/Lyft paradoxes)
- We must calculate conditional probabilities (within each nest)
  - $P_n(k \mid C_{mn}) = P_{kmn} = \exp(V_k \div m) \div \sum_{l \text{ in } Cmn} \exp(V_l \div m)$
- Inclusive value: value of nest's maximum expected utility
  - $IV_m = m \times \ln[\sum_{l \text{ in } Cmn} \exp(V_l \div m)]$
- Upper level probabilities
  - $P_n(m \mid C_n) = P_{mn} = \exp(IV_m) \div \sum_{l \text{ in } Cn} \exp(IV_l)$

EST. 1999
CIVIL
Engineering
MIST
Technology for Advancement
Military Institute of Science and Technology

Nested logit model

Alternative	Time	Cost
Car	5 minutes	$5.00
Bus 1 (blue)	15 minutes	$2.00
Bus 2 (red)	15 minutes	$2.00

$$V_k = (-0.2) \times \text{time}_k + (-1) \times \text{cost}_k$$

Buses 1 & 2 in nest with $m = 0.15$ (correlation = 0.85)

- Probabilities of choosing each alternative
  - $IV_{\text{bus}} = m_{\text{bus}} \times \ln[\exp(V_{\text{bus1}} \div m_{\text{bus}}) + \exp(V_{\text{bus2}} \div m_{\text{bus}})]\}]$
  - → $= 0.15 \times \ln[\exp(-5 \div 0.15) + \exp(-5 \div 0.15)]\}]$
  - → $= 0.15 \times \ln[3.34\text{E-}15 + 3.34\text{E-}15] = 0.15 \times -32.64 = -4.896$
  - $\text{Prob(car)} = \exp(V_{\text{car}}) \div [\exp(V_{\text{car}}) + \exp(IV_{\text{bus}})]$
  - → $= 0.00248 \div [0.00248 + 0.00748] = 0.00248 \div 0.00996 = 0.249$
  - $\text{Prob(buses)} = \exp(IV_{\text{bus}}) \div [\exp(V_{\text{car}}) + \exp(IV_{\text{bus}})] = \ldots = 0.751$

EST. 1999 CIVIL Engineering
MIST Technology for Advancement
Military Institute of Science and Technology

Types of independent variables

- Variables that vary across alternatives
  - Level-of-service variables: travel time, cost, crowding, comfort, etc.
  - → Can be included in every alternative's utility equation
  - → Can have generic (same for all alternatives) or alternative-specific coefficients
- Variables that do not vary across alternatives
  - Traveler characteristics: income, attitudes, etc.
  - → Can be included in every alternative's utility equation EXCEPT one (base)
  - → Must have alternative-specific coefficients
  - → Can be interacted with variables that vary across alternatives