# Discrete Choice Models

- Set of potential products $N = \{1, 2, \ldots, n\}$
- $\{0\} \equiv$ 'Outside option' (no purchase)
- $\boxed{\text{Choice model}}$: for all $S \subseteq N$, we have a distribution $\{\pi_j(S)\}_{j \in S \cup \{0\}}$
  - $\pi_j(S) = \mathbb{P}[\text{Customer picks item } j \in S]$ randomness over customers
    $\pi_0(S) = 1 - \sum_{j \in S} \pi_j(S) := 1 - \pi(S) = \mathbb{P}[\text{No purchase in } S]$
  - $\bar{S} \equiv N \setminus S$ (Complement set of items)
- Want choice models with less parameters (parsimonious)
- $\boxed{\text{Assortment optimization}}$
  - $P_j \equiv$ profit from item $j$ (exogenous)
    $R(S) = \text{Revenue from assortment } S = \sum_{j \in S} P_j \pi_j(S)$
  - $R^* = \max_{S \subseteq N} R(S)$

## * Independent demand model (perfect segmentation)

- $\pi_j(S) = \pi_j \quad \forall \; S : j \in S$
- Given non-negative $v_j$, $j \in \{0\} \cup N$, $v(N) = \sum_{j \in N} v_j$

$$\pi_j(S) = \frac{v_j}{\sum_{k=0}^{n} v_k} = \frac{v_j}{v_0 + v(N)} \quad \forall \; S \subseteq N$$

- Ignores 'demand recapture'. Can cause spiral-down of prices
- No basis in utility fns

## * Luce's choice axioms (Luce '59)

- For any $S \subset T$, define $\pi_S(T) = \sum_{j \in S} \pi_j(T)$

$$\pi_{S^+}(T) = \sum_{j \in S \cup \{0\}} \pi_j(T) = \pi_S(T) + \pi_0(T)$$

- Luce's Choice Axioms (LCA)

1) If $\pi_i(\{i\}) \in (0,1) \; \forall i \in T$, then $\forall R \subset S \cup \{0\}$, $S \subset T$

$$\pi_R(T) = \pi_R(S)\, \pi_{S^+}(T)$$

2) If $\pi_i(\{i\}) = 0$ for some $i \in T$, then $\forall S \subset T$ s.t. $i \in S$

$$\pi_S(T) = \pi_{S \setminus \{i\}}(T - \{i\})$$

<u>Thm</u> - A choice model satisfies the LCA iff $\exists\, v_j \geq 0$ s.t $\forall\, S \subseteq N$, $\forall\, j \in S$

$$\pi_j(S) = \frac{v_j}{\underbrace{\sum_{i \in S} v_i}_{v(S) = \sum_{i \in S} v_i} + v_0} = \frac{v_j}{v_0 + v(S)} \quad \circledast$$

Pf — 1) If $\pi_j(S) = \dfrac{v_j}{v_0 + v(S)}$ $\forall S$, then for any $R \subset S \cup \{0\}$, $S \subset T$

we have $\pi_R(S) = \dfrac{\sum_{j \in R} v_j}{v_0 + v(S)} = \left(\dfrac{v(R)}{v_0 + v(T)}\right) \cdot \left(\dfrac{v_0 + v(T)}{v_0 + v(S)}\right)$

$$= \pi_R(T) \Big/ \pi_{S^+}(T)$$

Also $\pi_i(\{i\}) = 0 \Rightarrow v_i = 0 \Rightarrow \pi_S(T) = \pi_{S - \{i\}}(T - \{i\}) = \dfrac{\sum_{j \in S,\, j \neq i} v_j}{v_0 + \sum_{j \in T,\, j \neq i} v_j}$

2) Suppose given choice model satisfies LCA

Then for $R = \{i\} \subset S \subset T = N$, we have

$$\pi_i(S) = \frac{\pi_i(N)}{\pi_{S^+}(N)} = \frac{\pi_i(N)}{\pi_0(N) + \sum_{j \in S} \pi_j(N)}$$

Now let $v_j = \pi_j(N)$ $\forall j \in N$, $v_0 = \pi_0(N)$

$$\Rightarrow \pi_j(S) = \frac{v_j}{\sum v_i + v_0}$$

# Random Utility Model for Choice

- (*) also arises via a random utility formulation
  - Each customer associates a <u>random</u> utility $U_j$ with product $j$, $U_0$ with no purchase
  - Given subset $S \subseteq N$, customer chooses $j \in S$ wp

$$\pi_j(S) = \mathbb{P}\left[U_j \geq \max_{i \in S \cup \{0\}} U_i\right]$$

  - Now suppose $U_j = \underbrace{u_j}_{\text{deterministic mean utility}} + \underbrace{\varepsilon_j}_{\text{iid r.v., 0 mean}}$

    - If $\varepsilon_j \sim N(0,1) \Rightarrow$ 'Probit' model
      - this has no simple closed-form like (*)
    - However, if $\varepsilon_j \sim \text{Gumbel}(\underset{\text{location}}{\cdot}, \underset{\text{scale}}{\cdot})$ with 0 mean

$$\text{then } \pi_j(S) = \frac{e^{\mu_j \phi}}{1 + \sum_{i \in S} e^{\phi \mu_i}} \quad \forall j \in S$$

(Multinomial Logit)

- Gumbel distribution - 2 parameters - location $\nu$, scale $\phi$

  - $F(x \mid \nu, \phi) = \exp\left[-\exp(-\phi(x-\nu))\right] \quad \forall x \in \mathbb{R}$

  - $f(x \mid \nu, \phi) = \phi \exp(-\phi(x-\nu)) \exp\left[-\exp(-\phi(x-\nu))\right]$

  - Euler constant

    $\text{Mean} = \nu + \dfrac{\gamma}{\phi}$ $\quad \left(\text{So } \nu = -\dfrac{\gamma}{\phi} \Rightarrow 0 \text{ mean}\right)$

    $\text{Mode} = \nu$, $\text{Median} = \nu - \ln(\ln(2))/\phi$, $\text{Variance} = \pi^2/6\phi^2$

  - Consider $X_1 = \mu_1 + \varepsilon_1$, $X_2 = \mu_2 + \varepsilon_2$, $\varepsilon_1, \varepsilon_2 \sim \text{Gumbel}\left(\frac{-\gamma}{\phi}, \phi\right)$

$$\mathbb{P}[X_1 \geq X_2] = \int_{-\infty}^{\infty} \mathbb{P}[x \geq X_2]\, f_{X_1}(x)\, dx$$

$$= \int_{-\infty}^{\infty} \underbrace{\phi e^{-\phi(x+\gamma/\phi-\mu_1)} e^{-e^{-\phi(x+\gamma/\phi-\mu_1)}}}_{f_{\varepsilon_1}(x-\mu_1)} \underbrace{e^{-e^{-\phi(x+\gamma/\phi-\mu_2)}}}_{F_{\varepsilon_2}(x-\mu_2)} dx$$

(Let $Z = e^{\phi\mu_1} + e^{\phi\mu_2}$)

$$= \frac{e^{\phi\mu_1}}{Z} \int_{-\infty}^{\infty} \phi Z e^{-\phi(x+\gamma/\phi)} e^{-Ze^{-\phi(x+\gamma/\phi)}} dx$$

$$= \frac{e^{\phi\mu_1}}{Z} \int_{-\infty}^{\infty} d\left(e^{-Ze^{-\phi(x+\gamma/\phi)}}\right) = \frac{e^{\phi\mu_1}}{e^{\phi\mu_1} + e^{\phi\mu_2}}$$

- The above calculation extends for multiple utilities

$$U_i = \mu_i + \varepsilon_i, \quad \varepsilon_i \sim \text{Gumbel}(-\gamma/\phi, \phi)$$

$$U_0 \equiv \varepsilon_0 \quad - \text{ 'Utility of outside option'}$$

$$\Rightarrow \mathbb{P}\left[U_i > \max\{U_0, \{U_j\}_{\substack{j\neq i \\ j\in S}}\}\right] = \frac{e^{\phi\mu_i}}{1+\sum_{j\in S} e^{\phi\mu_j}}$$

$$\Rightarrow \pi_j(S) = \frac{e^{\phi\mu_i}}{1+\sum_{j\in S} e^{\phi\mu_j}} \quad \forall S \subseteq N$$

---

## Independence of Irrelevant Alternatives

- Although the MNL model is parsimonious, it has one undesirable pathology (more generally, any model satisfying ✱)

- Suppose $\pi_j(S) = \dfrac{v_j}{v_0 + v(S)}$ $\Rightarrow$ $\boxed{\dfrac{\pi_j(S)}{\pi_j(S\cup\{k\})} = \dfrac{\pi_i(S)}{\pi_i(S\cup\{k\})} \quad \forall i, j \in S,\ k \notin S}$

This property is called the independence of irrelevant alternatives (IIA)

# * Negative consequence of IIA (Red bus-blue bus paradox)

- $S_0 = \{\text{red bus, car}\}$

  $S_1 = \{\text{red bus, blue bus, car}\}$

- $v_c, v_{rb}, v_{bb}$ are associated attraction values

  $v_{rb} = v_{bb}$ (utility independent of color of bus)

- $\Pi_{\{c\}}(S_0) = \dfrac{v_c}{v_{rb} + v_c}$ , $\Pi_{\{c\}}(S_1) = \dfrac{v_c}{v_{rb} + v_{bb} + v_c}$

$\Rightarrow$ Adding blue bus to S led to decrease in $\Pi_c$!

Problem - Ignores substitutability of products

# * Nested Logit Model - One fix for IIA

- MNL over clusters of 'substitutable' products
- $M = \{1, 2, \ldots, m\} \equiv$ Set of product clusters

  $N_i \equiv$ Set of products in cluster $i$, $S_i \equiv$ Subset offered
- Within cluster $i \equiv q_{j|i}(S_i) = \dfrac{v_{ij}}{V_i(S_i)} \equiv \mathbb{P}[\text{select item } j \in S_i]$
- Cluster selection $Q_i(S_1, S_2, \ldots, S_m) = \dfrac{(V_i(S_i))^{\gamma_i}}{v_0 + \sum_{j \in M} (V_j(S_j))^{\gamma_j}}$

  ($\gamma_i \in [0,1] \equiv$ dissimilarity in cluster $i$)

# * Universal Approximator Models

- MNL, Nested Logit are good models because they have 'easy' assortment optimization algos.

- However, many discrete choice models are far from MNL / NL

- We now consider more general models which act as 'universal approximators' for any discrete choice model

## 1) Mixture of Logit Models (McFadden & Train (2000))

$$\Pi_j(s) = \sum_{g \in G} \alpha^g \frac{v_j^g}{v_0^g + v^g(s)}$$

where $G \equiv$ Set of 'consumer types' / choice models

$$\sum_{g \in G} \alpha^g = 1$$

- ~~Mixture~~ of Mixed MNL can approximate any ~~distribution~~ discrete choice model arising from random utilities

- Difficult to do assortment optimization

## 2) Markov Chain Choice Model (Blanchet et al. (2013))

- General discrete choice model $\equiv$ Probability distrib$^{n}$ on 'preference lists' - permutations of $N \cup \{0\}$

  - Permutation $\sigma$ has prob $p(\sigma)$, $\sum_{\sigma} p(\sigma) = 1$

  - $\lambda_i := \pi_i(N) = \sum_{\sigma} p(\sigma)\, \mathbb{1}[\sigma(1)=i] = \sum_{\sigma} \mathbb{P}[\sigma(1)=i]$

    $\lambda_i \equiv$ 'first choice probabilities', $\sum_{i \in N \cup \{0\}} \lambda_i = 1$

  - If $\sigma(1)$ is not available, consumers switch to $\sigma(2)$

    - $\rho_{ij} = \mathbb{P}[\sigma(2)=j \mid \sigma(1)=i] \quad \forall\, i \neq j,\ i \in N,\ j \in N_+$

      Note - $\sum_{j \in N^+} \rho_{ij} = 1 \quad \forall\, i \in N$

    - $$\rho_{ij} = \frac{\pi_j(N \setminus \{i\}) - \pi_j(N)}{\pi_i(N)} \quad \forall\, i \neq j,\ i \in N,\ j \in N_+$$

      (can be used to estimate $\rho_{ij}$)

  - Eg - MNL with parameters $v_j = e^{\phi \mu_j}$ s.t. $v_0 + v(N) = 1$

    $$\Rightarrow \quad \lambda_i = v_i, \quad \rho_{ij} = \frac{v_j}{1 - v_i}$$

  - Eg - Mixed MNL - $\lambda_i = \sum_{g \in G} \alpha^g v_i^g$, $\rho_{ij} = \sum_{g \in G} \alpha^{g|i} \rho_{ij}^g$ where $\alpha^{g|i} = \frac{\alpha^g v_i^g}{[illegible]}$

    ($v_0^g + v^g(N) = 1\ \forall g$)

- Now given $\lambda_i \; \forall i \in N_+$, $\rho_{ij} \; \forall i \in N, j \in N_+$, we can set up the MC choice model as follows

  - For general choice model

$$\pi_i(S) = \mathbb{P}\left[ i \text{ before } j \text{ in } \sigma \; \forall j \in S \cup \{0\} \right]$$
$$= \sum_{\sigma} p(\sigma) \, \mathbb{1}\{ i \prec j \; \forall j \in S^+ \}$$

  - Instead of this, we assume customers sequentially look for products according to $\Lambda = \{\lambda_i\}$, $P = \{\rho_{ij}\}$ i.e., they sample the first product from $\Lambda$, and then transition according to $P$ till they find available product, or leave.

$$\Rightarrow \quad \pi_j(S) = \lambda_j + \sum_{i \in \bar{S}} \phi_i(S) \rho_{ij} \quad \forall j \in S$$
$$\phi_j(S) = \lambda_j + \sum_{i \in \bar{S}} \phi_i(S) \rho_{ij} \quad \forall j \in \bar{S}$$

<u>Note</u> - Random walk on $N_+$, with absorption in $S_+$

- MC choice model is a universal approximator AND has an 'easy' assortment optimization algorithm.