Discrete Choice Models of Demand With Consumer Individual Data
Lecture #7

Miguel I. Gómez
Dyson School of Applied Economics and Management
Cornell University

Overview

If you were in a supermarket buying a 12 pack of eggs and these were the only ones for sale which would you choose?

	Brand A	Brand B	Brand C
	Average Egg Weighs 55g	Average Egg Weighs 70g	Average Egg Weighs 60g
		Antibiotic and hormone free	Antibiotic and hormone free
			10% of Revenue donated to RSPCA
	Free Range	Free Range	Barn Raised
	All eggs appear the same	Some eggs appear different (e.g. Shell Colour)	All eggs appear the same
	Fed only on vegetables	Fed on grain and fish (high in Omega)	
	$4.90	$4.80	$3.00
Which would you choose?			

Source: http://mktresearch.org/wiki/Discrete_Choice_Experiments

Overview

- In the AIDS and Rotterdam models the consumer is somehow "representative" of the average consumer in the economy
- Assumed this consumer buys a little bit of every good in the system
- Models of "discrete choice" explain the observation that people usually purchase only one specific brand or product when faced with a choice of many differentiated items
- For most applications, the discrete choice model is a more accurate description of the decision-making process

Objective

- To learn the theoretical motivation and structure of discrete choice demand models

- To apply the model in situations where individual choice data are available

- To derive simple logit and mixed logit models and apply them to real-world data

- To motivate most complex simulation-based models

Discrete Choice / Random Utility Models

- Representative consumer assumption, not a good one for many differentiated products

- Assume consumers buy only one alternative or brand at a time – the one that has higher utility than all other choices

- This is the discrete choice model, conceptually preferred to the representative consumer approach

- Examples in our discipline?

Two Types of Discrete Choice Models Estimation

- **Simple logit -** If the consumer makes a straightforward choice among several brands
- **Mixed or random coefficients logit -** If the items to be chosen are highly differentiated and the choosers have a number of attributes (e.g., age, education, race, etc.) that are likely to be important in their choice, and correlated with the attributes they choose

Discrete Choice Model – Choice Set

- What do we need to define?
  - Choice set that includes:
    - Mutually exclusive,
    - Exhaustive, and
    - Finite number of alternatives
- Examples
  - Suppose you want to know consumers' choice of carbonated soft drink (CSD) in brand level. In this case, your choice set should be: $\{Coke, Pepsi, Sprite, Fanta, Others\}$
  - How do you define your choice set if you want to investigate the effect of calorie content?
  - How do you define your choice set if you want to know the effect of package size?
  - How do you define your choice set if you are interested in commuters' mode choice?

Discrete Choice Model – Introduction (Cont.)

- Consider the choice situation in which J alternatives are available and alternative 2 is chosen by the decision maker.

Alternative	A_1	A_2	$\cdots$	A_J
Choice probability	P_1	P_2	$\cdots$	P_J
Choice		X		

- Because alternative 2 is chosen, the likelihood of alternative 2, P_2 must be maximized.
- The parameterization of the likelihood i.e. $P_j(X, \beta)$ where X is the set of variables and β is the set of parameters, allows us to estimate the effect of the variables on the decision makers' choice.
- We can estimate β by maximizing the likelihood of the alterantive chosen by the decision maker.

Discrete Choice Model – Introduction (Cont.)

- Setup
  - Decision maker: $n = 1,2,\cdots,N$
  - Alternative: $j = 1,2,\cdots,J$
  - Utility: $U_{nj} = V_{nj} + \varepsilon_{nj} = X_{nj}\beta + \varepsilon_{nj}$ (1)

Random part

Deterministic part

- This approach is also called a "random utility model" because consumer's tastes are represented by adding the random error term ε_{nj} to the utility function.
- The random error ε_{nj} reflects the fact that decision makers are going to prefer different alternatives given that they prefer the attributes of each alternative differently.
- The deterministic part, V_{nj} is parameterized by β, which allows us to estimate the effect of the variable, X_{nj} on utility.

Discrete Choice Model – Introduction (Cont.)

- We want to find the choice probability in each choice situation.
- The probability that decision maker n selects alternative i is written as:

$$\begin{aligned} P_{ni} &= Prob\left(U_{ni} > U_{nj}, \forall j \neq i\right) \\ &= Prob\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall j \neq i\right) \\ &= Prob\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall j \neq i\right) \\ &= \int I\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall j \neq i\right) f(\varepsilon_n) d\varepsilon_n \quad (2) \end{aligned}$$

- How can we obtain the choice probability, P_{ni}?
- We need to make a parametric assumption on ε_{nj}.

Discrete Choice Model – Introduction (Cont.)

- If we assume that ε_{nj} is iid extreme value (EV), the model is called logit model.
- By using the EV assumption, we can derive a closed form formula for the choice probability P_{ni}, which makes the estimation easy.
- The assumption implies that the unobserved factors are uncorrelated over alternatives, as well as having the same variance for all alternatives.
- Is that assumption valid? Perhaps, no.
- Example
  - Suppose you examine person's transport mode choice.
  - The choice set is defined by: $\{Bus, Train, Car, Bike\}$.
  - A person dislikes traveling by bus because of the presence of others. This person may make a similar reaction to train because there are other riders in the train as well.
  - The unobservable factors affecting bus and train are likely to be correlated rather than independent.
- The mixed logit model overcome this problem!

Discrete Choice Model – Logit Model

- The logit model is written as:

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \forall j, \varepsilon_{nj} \sim iid\ EV \quad (3)$$

- The property of EV
  - Density function: $f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}}$ (4)
  - Distribution function: $F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}}$ (5)
  - Expected value: $E(\varepsilon_{nj}) = \gamma$ (Euler's constant) (6)
  - Variance: $V(\varepsilon_{nj}) = \frac{\pi^2}{6}$ (7)
- If $\varepsilon_{nj} \sim iid\ EV$, the choice probability, P_{ni} is written as:

$$\boxed{P_{ni} = \frac{\exp(V_{ni})}{\sum_{j=1}^{J} exp(V_{nj})} \quad (8)}$$

Discrete Choice Model – Logit Model (Cont.)

Example

- You want to analyze household's choice of heating system: gas heating system, electric heating system, or oil-fuel heating system. Assume there is no other alternative and two heating systems cannot be used at the same time.
  - Why do we need the second sentence?
  - What is the choice set?

Discrete Choice Model – Logit Model (Cont.)

- Suppose that the utility from using each heating system depends only on the purchase price (PP_j) and the annual operating cost (OC_j) where $j = g, e, o$ represents each alternative.

  - Describe the logit choice probability for each alternative.
  - Suppose electric old system is chosen. What should happen?

Discrete Choice Model – Logit Model (Cont.)

Identification of logit model 1

- Consider the following logit model:

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \forall j, \varepsilon_{nj} \sim iid\ EV \quad (9)$$

- The probability that decision maker chooses an alternative i is given by:

$$P_{ni} = \frac{\exp(V_{ni})}{\sum_{j=1}^{J} exp(V_{nj})} \quad (10)$$

- What does happen when a constant ω is added to utility?

$$P_{ni} = \frac{\exp(V_{ni}+\omega)}{\sum_{j=1}^{J} exp(V_{nj}+\omega)} = \frac{\exp(V_{ni})exp(\omega)}{\sum_{j=1}^{J} exp(V_{nj})exp(\omega)} = \frac{\exp(V_{ni})}{\sum_{j=1}^{J} exp(V_{nj})} \quad (11)$$

- The probability does not change. This implies that only differences in utility matters, and all constant terms cannot be identified.

Discrete Choice Model – Logit Model (Cont.)

Example

- In a brand choice model, we often use a brand-specific intercept which measures brand equity or intrinsic preference. For example, suppose you want to analyze consumer n's choice of Coke and Pepsi each of which is described by price ($P_{n,j}$) and brand specific intercept (α_j). Then, you construct the following random utility model.

$$U_{n,Coke} = \alpha_{Coke} + \beta P_{n,Coke} + \varepsilon_{n,Coke} \quad (12)$$

$$U_{n,Pepsi} = \alpha_{Pepsi} + \beta P_{n,Pepsi} + \varepsilon_{n,Pepsi} \quad (13)$$

- To estimate the utility, either α_{Coke} or α_{Pepsi} must be normalized to zero. That is, the model should become:

$$U_{n,Coke} = \beta P_{n,Coke} + \varepsilon_{n,Coke} \quad (14)$$

$$U_{n,Pepsi} = \alpha_{Pepsi} + \beta P_{n,Pepsi} + \varepsilon_{n,Pepsi} \quad (15)$$

Discrete Choice Model – Logit Model (Cont.)

Identification of logit model 2

- Consider the following logit model:

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$
$$= \beta x_{nj} + \gamma_j x_n + \varepsilon_{nj}, \forall j, \varepsilon_{nj} \sim iid\ EV \quad (16)$$

- The probability that decision maker chooses alternative i is given by:

$$P_{ni} = \frac{\exp(\beta x_{ni} + \gamma_i x_n)}{\sum_{j=1}^{J} exp(\beta x_{nj} + \gamma_j x_n)} \quad (17)$$

- What does happen when a constant η is added to γ?

$$P_{ni} = \frac{\exp(\beta x_{ni} + (\gamma_i + \eta) x_n)}{\sum_{j=1}^{J} exp(\beta x_{nj} + (\gamma_j + \eta) x_n)} = \frac{\exp(\beta x_{ni} + \gamma_i x_n) exp(\eta x_n)}{\sum_{j=1}^{J} exp(\beta x_{nj} + \gamma_j x_n) exp(\eta x_n)} = \frac{\exp(\beta x_{ni} + \gamma_i x_n)}{\sum_{j=1}^{J} exp(\beta x_{nj} + \gamma_j x_n)} \quad (18)$$

- The probability does not change. This implies that we cannot estimate all the parameters associated with the variable that does not depend on the alternative.

Discrete Choice Model – Logit Model (Cont.)

Example

- Let's continue the CSD example. Now you believe that household income is an important choice determinant. However, you cannot estimate all the parameter estimates for the income variable, and need to normalize one of them to zero, so that your final model should be:

$$U_{n,Coke} = \beta P_{n,Coke} + \varepsilon_{n,Coke} \quad (19)$$

$$U_{n,Pepsi} = \alpha_{Pepsi} + \beta P_{n,Pepsi} + \gamma INCOME + \varepsilon_{n,Pepsi} \quad (20)$$

Discrete Choice Model – Logit Model (Cont.)

- Another solution for these problems
- Incorporate outside option such as "non-purchase" or "others."
- We usually assume that the utility from outside option is just an error term which means the deterministic part for outside option is zero. This assumption enables you to estimate all the parameters for your choice alternatives except for the outside option.
- For example, if we include "other products" into the examples in slides 13 and 15, then the following model is estimable.

$$U_{n,Coke} = \alpha_{Coke} + \beta P_{n,Coke} + \gamma INCOME + \varepsilon_{n,Coke} \quad (21)$$

$$U_{n,Pepsi} = \alpha_{Pepsi} + \beta P_{n,Pepsi} + \gamma INCOME + \varepsilon_{n,Pepsi} \quad (22)$$

$$U_{n,Others} = \varepsilon_{n,Others} \quad (23)$$

Discrete Choice Model – Logit Model (Cont.)*

Estimation

- Idea
  - We observe decision maker's choice in the data. If one alternative is chosen, the probability that alternative is chosen must be maximized. By doing so, we can estimate the parameters that consist of the utility and choice probability.

- This is called the Maximum Likelihood method.

Discrete Choice Model – Logit Model (Cont.)*

- Consider the logit model:

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \forall j, \varepsilon_{nj} \sim iid\ EV \tag{24}$$

- Let the likelihood for decision maker n be L_n and the likelihood for all decision makers L.

- $L_n = \prod_{j=1}^{J} (P_{nj})^{y_{nj}}$ where y_{nj} takes 1 if decision maker n chooses alternative j, 0 otherwise. (25)

- $L = \prod_{n=1}^{N} L_n$ (26)

Discrete Choice Model – Logit Model (Cont.)*

- We can obtain parameter estimates in equation 24 by maximizing equation 26.

- We usually write the likelihood as $L(\beta)$ because it is a function of parameters β.

- In the maximization process, we take the natural log of $L(\beta)$, and the log of $L(\beta)$is called log-likelihood value. We donate it as $l(\beta)$.

Discrete Choice Model – Logit Model (Cont.)*

Model selection

1. Likelihood ratio test
2. Akaike Information Criterion (AIC)
3. Bayesian Information Criterion (BIC)

Discrete Choice Model – Logit Model (Cont.)*

1. **Likelihood ratio test**

- Compare the likelihood of a null model with an alternative model.
- The test statistic is: $LR = -2log\frac{L(\widehat{\beta_0})}{L(\hat{\beta})} = -2\{l(\widehat{\beta_0}) - l(\hat{\beta})\}$ (27)
- $LR \sim \chi^2(g)$, where g is the difference in the number of parameters in the null and alternative model
- If LR exceeds a critical value, then the null model is rejected.
  - $L(\widehat{\beta_0})$: likelihood value for a null model
  - $L(\hat{\beta})$: likelihood value for an alternative model
  - $l(\widehat{\beta_0})$: log-likelihood value for a null model
  - $l(\hat{\beta})$: log-likelihood value for an alternative model

Discrete Choice Model – Logit Model (Cont.)*

Example

- Let's recall the example in slide 15. At that time, you decided to include the income variable. So, the model without and with income variables must statistically be compared.
- Null model: model without the income variable
- Alternative model: model with the income variable
- We can calculate LR using the log-likelihood value from the null and alternative models.
- In this case, $g = 1$.
- If LR is above the critical value, then we can reject the null model, and conclude the model with the income variable is better than the other.

Discrete Choice Model – Logit Model (Cont.)*

2. Akaike Information Criterion (AIC)

$AIC = -2l(\hat{\beta}) + 2k$, where k is the number of parameters (28)

3. Bayesian Information Criterion (BIC)

$BIC = -2l(\hat{\beta}) + k \times logN$, where k is the number of parameters and N is the number of observations (29)

- The smaller AIC and BIC are better.
- BIC penalizes more the increased number of parameters.

Discrete Choice Model – Logit Model (Cont.)

Example 1

- Suppose you work for a marketing research company, and your client in the milk industry ask you to investigate the effect of in-store advertising.
- The data you received contains:
  - Choice of 4 milk brands by each consumer: Bärenmarke, Landliebe, Weihenstephan, and Berchtesgadener Land
  - Price and in-store advertising for those brands
- There are 10 choice situations for each consumer.
- For details, see the data file (Milk_Brands_example.xlsx)
- How do you specify the utility?
- What is the choice probability for each milk brand?

STATA Syntax

```
* Data loading
clear
set more off
import excel
"C:\Users\mig7\Box\My_Documents\AEM6700\AEM6700_2020\AEM6700_Lectures_Fall2020\Milk_Brands_DATA.xlsx",
sheet("Milk Choice") firstrow
* Change the data format from wide to long
reshape long price prom choice, i(sn) j(brand)
* Create alternative specific binary variables
gen brand1 = (brand == 1)
gen brand2 = (brand == 2)
gen brand3 = (brand == 3)
gen brand4 = (brand == 4)

asclogit choice brand1 brand2 brand3 price prom, case(sn) alternatives(brand) noconstant
estat ic
```

Discrete Choice Model – Logit Model (Cont.)

- Utility for consumer n from choosing a milk brand j at situation t is written as:

$$U_{njt} = \alpha_j \ + \beta PRICE_{njt} \ + \gamma INSTORE\_AD_{njt} \ + \varepsilon_{njt} \quad (30)$$

Table 1: Estimation Result of the Milk Choice Model

Variable	Estimate	t-ratio
Price	-4.903	-21.23
In-store advertising	2.15	27.02
Bärenmarke	1.959	21.93
Landliebe	0.978	11.33
Weihenstephan	-0.108	-1.25
Log likelihood value	-1990.572	
AIC	$3,991.14$	
BIC	$4,019.14$	

- Price coefficient is negative as expected.
- In-store advertising has positive effect on utility.
- Bärenmarke and Landliebe are preferred than Berchtesgadener Land.

Discrete Choice Model – Logit Model (Cont.)

Example 2

- Suppose you work for the same marketing research company, and your client in the milk industry ask you to investigate the effect of in-store advertising.
- **But** instead the data you received contains:
  - Choice of 4 milk brands by each consumer: Bärenmarke, Landliebe, Weihenstephan, and **All Others**
  - Price and in-store advertising for the three brands
- There are 10 choice situations for each consumer.
- For details, see the data file (**Milk_Brands_example2**.xlsx)
- How do you specify the utility?
- What is the choice probability for each milk brand?

STATA Syntax

```
* Data loading
clear
set more off
import excel
"C:\Users\mig7\Box\My_Documents\AEM6700\AEM6700_2020\AEM6700_Lectures_Fall2020\MILK_Brands_Example2.xls
x", sheet("L7_DATA1") firstrow

* price4 and promo4 are set to be zero because deterministic part of the utility from choosing outside option is assumed to
be zero.
gen price4 = 0
gen promo4 = 0

* Change the data format from wide to long
reshape long price promo choice, i(sn) j(alt)

* Create alternative specific binary variables
gen alt1 = (alt == 1)
gen alt2 = (alt == 2)
gen alt3 = (alt == 3)

* Model 1: Simple logit
asclogit choice alt1 alt2 alt3 promo price, case(sn) alternatives(alt) noconstant
estat ic
```

Discrete Choice Model – Logit Model (Cont.)

- Utility for consumer n from choosing a milk brand j at situation t is written as:

$$U_{njt} = \alpha_j \;\; + \beta PRICE_{njt} \; + \gamma INSTORE\_AD_{njt} \; + \varepsilon_{njt} \quad (30)$$

```
                                                Wald chi2(5)     =    1088.53
Log likelihood = -3500.0291                     Prob > chi2      =     0.0000
```

choice	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
alt						
alt1	.4794507	.1966398	2.44	0.015	.0940437	.8648577
alt2	.9223535	.1805285	5.11	0.000	.5685242	1.276183
alt3	1.460611	.187888	7.77	0.000	1.092357	1.828865
promo	2.058831	.0691533	29.77	0.000	1.923293	2.194369
price	-.907436	.1980365	-4.58	0.000	-1.29558	-.5192915

```
. estat ic
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	12,000	.	-3500.029	5	7010.058	7047.022

Discrete Choice Model Own Price Elasticity

Recall that

$$P_i = \frac{\exp(V_i)}{\sum_{j=1}^{J} exp(V_j)}$$

P_i is the probability that brand *i* is chosen; or the predicted market share of brand *i*.

Let us call $P_i = S_i$, the predicted market share of brand *i.*

Our model for milk brands is:

$U_j = \hat{\alpha}_j + \hat{\beta}\ PRICE_j + \hat{\gamma}\ INSTORE\_AD_j + \varepsilon_j$ for *j = 1, 2, 3, 4* (the four brands)

And the deterministic component of the utility is

$V_j = \hat{\alpha}_j + \hat{\beta}\ PRICE_j + \hat{\gamma}\ INSTORE\_AD_j$

Discrete Choice Model Elasticities

- Own-price elasticities (Call variable *PRICE = P)*
  - Share derivative is:

$$\frac{\partial S_i}{\partial p_i} = \beta exp(V_i)\left\{\sum_{j=1}^{J} exp(V_j)\right\}^{-1} - \beta exp(V_i) exp(V_i)\left\{\sum_{j=1}^{J} exp(V_j)\right\}^{-2}$$

$$= \beta S_i - \beta {S_i}^2 = \beta S_i(1 - S_i)$$

  - So, the own-price elasticity is:

$$E_{ii} = \frac{\partial S_i}{\partial P_i}\frac{P_i}{S_i} = \beta P_i(1 - S_i)$$

- P_i comes form the data, and β and S_i are estimated from the model

Discrete Choice Model: Cross Price Elasticities

- Share derivative is:

$$\frac{\partial S_i}{\partial P_j} = -exp(V_i)(\beta)exp\left(V_j\right)\left\{\sum_{j=1}^{J} exp\left(V_j\right)\right\}^{-2} = -\beta S_j S_i$$

- So, the cross-price elasticity is:

$$E_{ij} = \frac{\partial S_i}{\partial P_j}\frac{P_j}{S_i} = -\beta P_j S_j$$

- P_j comes form the data, and β and S_j are estimated from the model

Discrete Choice Model – Logit Model (Cont.)

- The power of the logit model is evident – it is a simple way of describing utility maximizing behavior when the choice environment may consist of dozens of products that differ in price, attributes and location. However, it also has some very important limitations. The notable limitations include:

1. Taste variation
2. Unrealistic substitution pattern

Discrete Choice Model – Logit Model (Cont.)

1. Taste variation

- By including demographic attributes such as age, education and race, the logit model can explain systematic variation in tastes that depend upon the type of individuals involved (e.g., poor people tend to buy more hamburgers than rich people).
- But it is not well suited for describing random variations in taste, or decisions that are driven by idiosyncratic factors, this is also known as "unobserved heterogeneity."
- If the variation in tastes among decision makers is dependent on observed factors and some random element interacted with those factors, then the error term for the entire model depends on the observed attributes for either the decision maker or the choice so cannot possibly be *iid*.
- This is a violation of the assumptions that underlie the logit model.

Discrete Choice Model – Logit Model (Cont.)

2. Unrealistic substitution pattern

- The probability of choosing all products must clearly sum to 1.0
- Therefore, when prices change, or some other attribute of the product changes, the probability of choosing one product either rises (price falls) or falls (price rises) and the probability of choosing other products must necessarily change in response.
- The logit model implies a very strict pattern of substitution among products, which may or may not be a good thing, called the "independence of irrelevant alternatives (IIA)" property.

Discrete Choice Model – Logit Model (Cont.)

2. Unrealistic substitution pattern

- IIA property: The ratio of probabilities between two choices is independent of the utility obtained from any other choice.
- If this makes sense in the particular situation at hand, it is a good thing, but often it does not.
- Mathematically, IIA implies that the ratio of probabilities of two alternatives does not depend upon the utility obtained from a third.
- Consider the logit model: $U_{nj} = V_{nj} + \varepsilon_{nj}, \forall j, \varepsilon_{nj} \sim iid\ EV$ (31)
- The ratio of the probabilities to choose alternative 1 and 2 are:

$$\frac{P_{n1}}{P_{n2}} = \frac{\exp(V_{n1})/\sum exp(V_{nj})}{\exp(V_{n2})/\sum exp(V_{nj})} = \frac{\exp(V_{n1})}{\exp(V_{n2})} = exp(V_{n1} - V_{n2}), \quad (32)$$

which does not depend on other alternatives.

Discrete Choice Model – Logit Model (Cont.)

2. Unrealistic substitution pattern

- To see how this problem arises in a more concrete way, consider the famous “red bus / blue bus” problem.
- Suppose a commuter can travel either by train or by a red car and her probability of choosing each is 1/2. Now, suppose we introduce the option of traveling by a blue car.
- Because the logit model implies that the ratio of probabilities must stay the same, the probability of choosing train falls to 1/3, and the probability of red car falls to 1/3 as well to accommodate the probability of traveling by blue car (the probability to choose a blue car is 1/3).
- This is clearly unrealistic as the probability of choosing the train should not fall if another type of car, which provides the same attributes as the other car, is introduced. It should be the case that the probability of choosing blue car and the probability of choosing red car is 1/4 respectively, and the probability of choosing train is 1/2.
- But this would violate the constant-odds ratio property. If a logit model is being used to predict demand, it will under-predict the demand for train ridership and over-predict the demand for each type of car.

Discrete Choice Model – Logit Model (Cont.)

Suppose that individuals have the choice out of three restaurants, Chez Panisse (C), Lalime's (L), and the Bongo Burger (B). Suppose we have two characteristics, price and quality:

price $P_C = 95$, $P_L = 80$, $P_B = 5$,

quality $Q_C = 10$, $Q_L = 9$, $Q_B = 2$

market share $S_C = 0.10$, $S_L = 0.25$, $S_B = 0.65$.

These numbers are roughly consistent with a logit model where the utility associated with individual *i* and restaurant *j* is

$$U_{nj} = -0.2 \cdot P_j + 2 \cdot Q_j + \varepsilon_{nj},$$

Discrete Choice Model – Logit Model (Cont.)

Now suppose that we raise the price at Lalime's to 1000 (or raise it to infinity, corresponding to taking it out of business).

The simple logit model predicts that the market shares for Lalime's gets divided by Chez Panisse and the Bongo Burger, proportional to their original market share, and thus $\tilde{S}_C$ = 0.13 and $\tilde{S}_B$ = 0.87: most of the individuals who would have gone to Lalime's will now dine (if that is the right term) at the Bongo Burger.

That seems implausible. The people who were planning to go to Lalime's would appear to be more likely to go to Chez Panisse if Lalime's is closed than to go to the Bongo Burger, implying $\tilde{S}_C \approx 0.35$ and $\tilde{S}_B \approx 0.65$.

Discrete Choice Model – Logit Model (Cont.)

2. Unrealistic substitution pattern

- Another way of seeing this problem is to consider the logit cross-price elasticity.
- Earlier, we derived the logit cross-price elasticity between alternative i and j as

$$E_{ji} = \frac{\partial S_j}{\partial P_i}\frac{P_i}{S_j} = \beta P_i S_i \quad (33)$$

- Clearly, this elasticity depends only on the price of the other product and its market share – nothing about product j enters this expression. Therefore, the cross-price elasticity is constant with respect to all other j products and changes in market share depend only on the existing share of the other product. This is unrealistic.

Choice Data – Primary Data

- A choice experiment or conjoint experiment allows us to collect choice data (to see in next classes).
- There are several choice situations in the choice experiment.
- In each choice situation, researchers let the subjects choose one alternative among several.
- Each choice alternative is defined by different attribute(s) and different attribute level(s).

Choice Data – Primary Data, Example

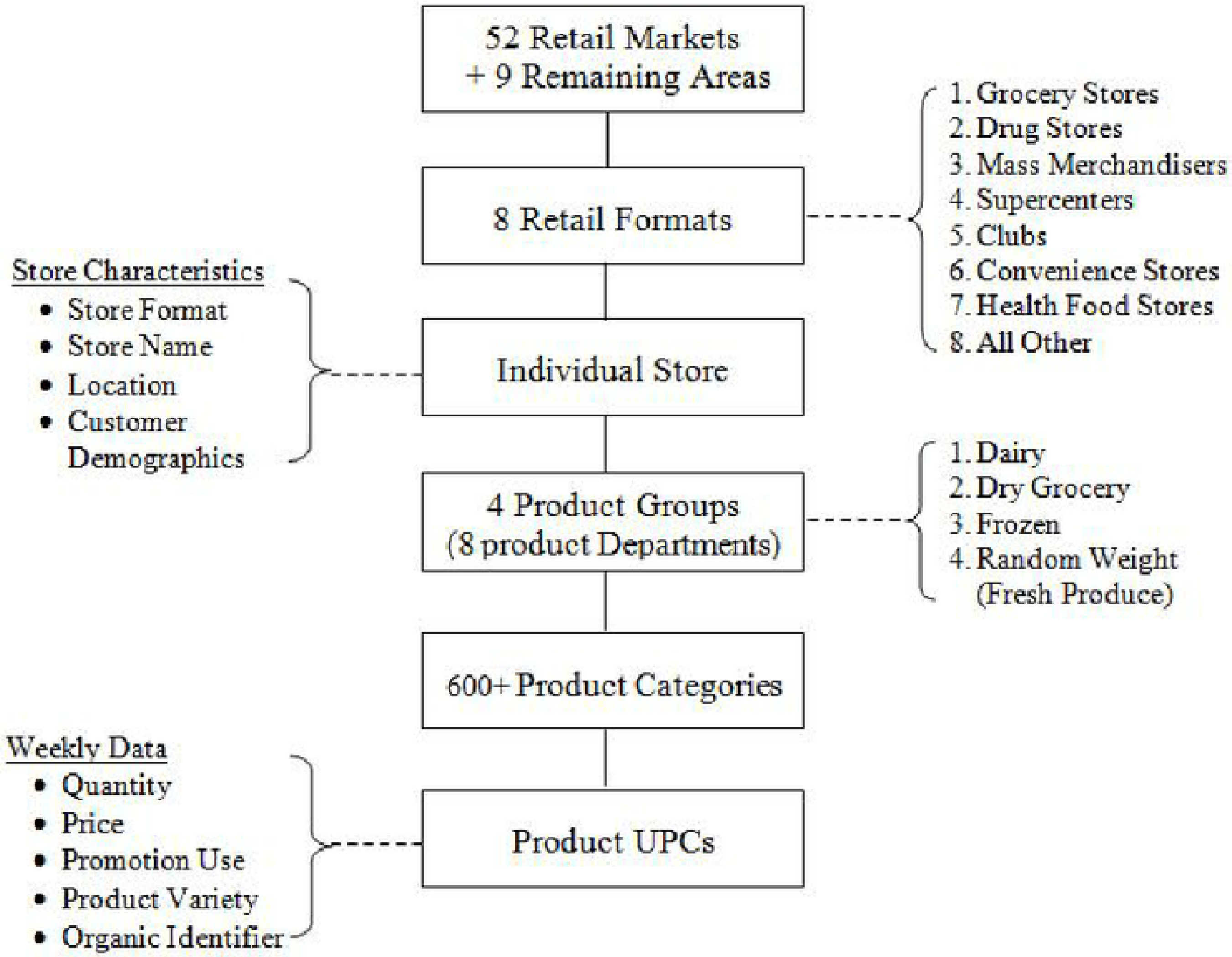
If you were in a supermarket buying a 12 pack of eggs and these were the only ones for sale which would you choose?

	Brand A	Brand B	Brand C
	Average Egg Weighs 55g	Average Egg Weighs 70g	Average Egg Weighs 60g
		Antibiotic and hormone free	Antibiotic and hormone free
			10% of Revenue donated to RSPCA
	Free Range	Free Range	Barn Raised
	All eggs appear the same	Some eggs appear different (e.g. Shell Colour)	All eggs appear the same
	Fed only on vegetables	Fed on grain and fish (high in Omega)	
	$4.90	$4.80	$3.00
Which would you choose?			

Choice Data – Secondary Data

- Household panel scanner data
- Panels of households are often assembled to record everything that they purchase in a given time frame - usually over a series of weeks extending up to two or three years.
- The household panel scanner data does not include prices for items a household may have looked at, but not purchased.
- Researchers either impute those prices by using transactions in similar choice situations (in terms of week, location, and store) or estimate prices using a hedonic price model.

Choice Data – Secondary Data, Example: Structure and Description of A.C Nielsen Homescan Data Set



References

- Kuhfeld, W. (2010). Experimental design and choice modeling macros. Retrieved from http://support.sas.com/resources/papers/tnote/tnote_marketresearch.html.
- Louviere, J. J., Hensher, D. A., & Swait, J. (2000). *Stated choice methods: Analysis and applications.* Cambridge, U.K.: Cambridge University Press.
- Malhotra, N. K. (2008). *Marketing research: An applied orientation*. Pearson.
- Meas, T., Hu, W., Batte, M. T., Woods, T. A., & Ernst, S. (2014). Substitutes or complements? Consumer preference for local and organic food attributes. *American Journal of Agricultural Economics*, aau108.
- Train, K. (2009). *Discrete choice methods with simulation.* Cambridge university press. Retrieved from http://eml.berkeley.edu/books/choice2.html.

Advanced Topics

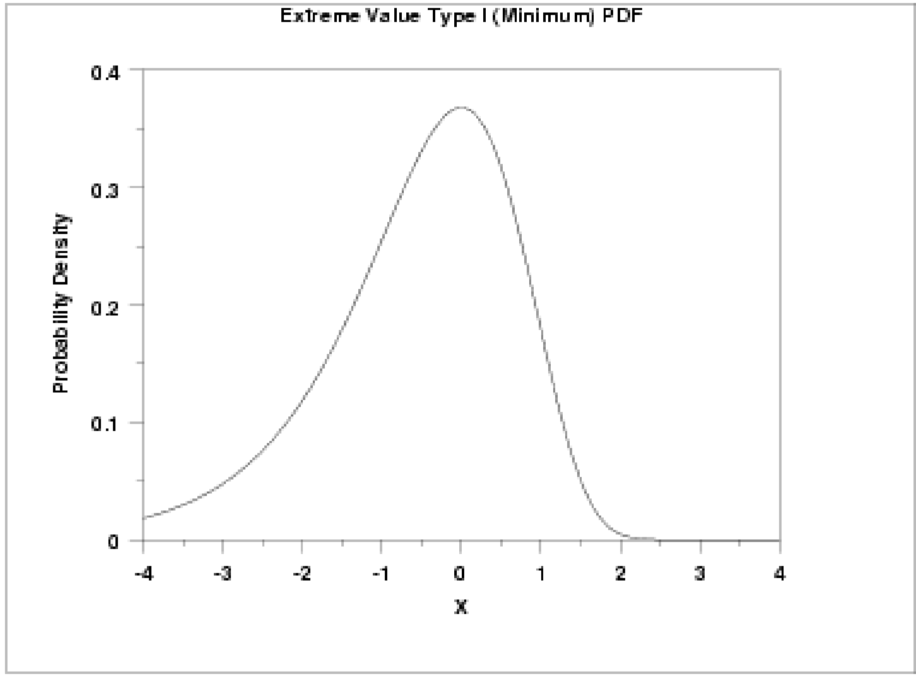
(not required for this course)

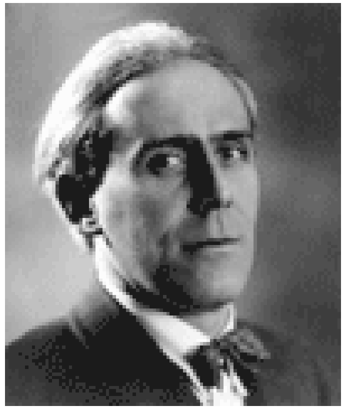
Type I Extreme Value Functions*

- The form of the model that is determined by the assumption on distribution of the random error

- If normally distributed, then the model becomes a probit

- Assuming the distribution of consumer heterogeneity is i.i.d. across products and consumers, Type I Extreme Value





Emil Gumbel

$$f(x) = \frac{1}{\sigma} \exp\left(-z - \exp(-z)\right)$$

where z=(x-μ)/σ, μ is the location parameter, and σ is the distribution scale (σ>0)