

# Discrete Choices and Extreme Value Distributions

Motoaki Takahashi

Graduate Urban Economics at the University of Osaka

November 6, 2025

## Review of the Rosen-Roback model and motivation

- ▶ We learned the Rosen-Roback model in the previous lecture.
- ▶ There, everyone's welfare is exactly equalized across all locations.
- ▶ But, you live in your location because you prefer your location to the other locations.
- ▶ And different people have different tastes over locations.
  - ▶ I love Osaka. Whatever happens, I will live in Osaka until my last day.
  - ▶ I will never move from Tokyo.
  - ▶ I like to live in a quiet, rural area.

## Discrete choices

- ▶ If we treat locations as discrete points, location choices are discrete choices.
- ▶ Discrete choices:
  - ▶ I choose 3 from  $\{1, 2, 3, 4, 5\}$ .
  - ▶ Feasible locations are  $\{\text{Toyonaka, Ikeda, Ibaraki}\}$ . I choose to live in Toyonaka.
  - ▶ I can buy goods from  $\{\text{China, Korea, Japan}\}$ . Eventually I decide to buy goods from Japan.
- ▶ Not discrete choices:
  - ▶ Locations are represented as points over  $[0, 1]$ . I choose to live in point 0.55425.
  - ▶ Let  $x$  be my consumption of beer. My utility function is

$$u(x) = -(x - 2)^2.$$

Maximizing this function, I choose to drink 2 liters of beer.

## Review of econometrics classes: Logit (1)

- ▶ Actually, if we consider only two locations, we do not need extreme value distributions.<sup>1</sup>
  - ▶ e.g. Toyonaka and the rest of Japan.
- ▶ Instead, all we need is logistic distributions.
- ▶ Consider the following simple two-location model.
- ▶ There are two locations:  $A$  and  $B$ .
- ▶ If an individual lives in  $B$ , she has deterministic utility  $v^B$ .
- ▶ If she lives in  $A$ , her utility is

$$v^A + \epsilon,$$

where  $v^A$  is the deterministic part of utility in  $A$  and  $\epsilon$  is the stochastic part.<sup>2</sup>

---

<sup>1</sup>It's ok to assume extreme value distributions, though.

<sup>2</sup>It is often called idiosyncratic preference shocks.

## Review of econometrics classes: Logit (2)

- ▶ The individual chooses location  $A$  if  $v^A + \epsilon > v^B$ .
- ▶ She chooses location  $B$  if  $v^B \geq v^A + \epsilon$ .
- ▶ Assume that the shock  $\epsilon$  follows the logistic distribution.
- ▶ The distribution function of the logistic distribution with scale parameter  $s$  and location parameter  $\mu$  is

$$F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{s}}}.$$

- ▶ For normalization, let  $\mu = 0$ .
- ▶ Then the probability that the individual chooses to live in  $B$  is

$$\begin{aligned} & P[v^A + \epsilon \leq v^B] \\ &= P[\epsilon \leq v^B - v^A] \\ &= F(v^B - v^A) \\ &= \frac{1}{1 + e^{-\frac{v^B - v^A}{s}}}. \end{aligned}$$

## Review of econometrics classes: Logit (3)

- ▶ Then the probability that the individual chooses to live in  $A$  is

$$1 - F(v^B - v^A) = \frac{e^{-\frac{v^B - v^A}{s}}}{1 + e^{-\frac{v^B - v^A}{s}}}.$$

- ▶ We considered only one individual's behavior.
- ▶ But, now suppose that there is a mass or continuum of such people.
- ▶ That is, a unit mass of people receive idiosyncratic shocks from the same logistic distribution independently.
- ▶ Then now the probabilities become *shares*.
- ▶ The share  $F(v^B - v^A)$  of these people choose to live in  $B$ .
- ▶ The share  $1 - F(v^B - v^A)$  of these people choose to live in  $A$ .

## From two locations to many (but discrete) locations

- ▶ The two-location example works well to illustrate basics of the discrete choice approach in urban/spatial economics.
- ▶ But, in the real world, there are many locations.
  - ▶ Osaka, Kyoto, Shiga, Wakayama, etc... in Japan.
  - ▶ Toyonaka, Ibaraki, Minoh, etc... in Osaka.
- ▶ Extreme value distributions are a powerful tool for modeling discrete choices among multiple alternatives.
- ▶ Assuming extreme value distributions, we can derive analytical expressions for shares (the population share of each location) and welfare.
  - ▶ These share and welfare formulae resemble formulae from the constant-elasticity-of-substitution (CES) function.
  - ▶ Anderson et al. (1987) for single layer cases.
  - ▶ See Verboven (1996) and appendix B of Berger et al. (2022) for multi layer cases.

# Setup

- ▶ There are  $n$  locations. Locations are indexed by  $i = 1, \dots, n$ .
- ▶ We consider only individuals' location choices.
  - ▶ We do not consider how wages, rents, and other prices and quantities are determined.
  - ▶ That is, here we do not consider a general (or even partial) equilibrium model.
  - ▶ We consider only individuals' optimization problem over locations.
- ▶ There is a unit mass of individuals.

# Utilities

- ▶ Pick up an infinitesimal individual from the unit mass.
- ▶ If she lives in location  $i \in \{1, \dots, n\}$ , her utility is

$$v_i \epsilon_i,$$

where

- ▶  $v_i > 0$  is the deterministic part of the utility (like amenity-adjusted real wages),
  - ▶  $\epsilon_i$  is the stochastic part of the utility.
- ▶  $\epsilon_i$  is often called the idiosyncratic preference or amenity shock.

## The Fréchet distribution

- ▶ We assume that across all locations and all infinitesimal individuals, the idiosyncratic preference shock  $\epsilon_i$  independently and identically follows the Fréchet distribution

$$F(x) = e^{-x^{-\theta}},$$

where  $\theta > 0$  is the shape parameter.

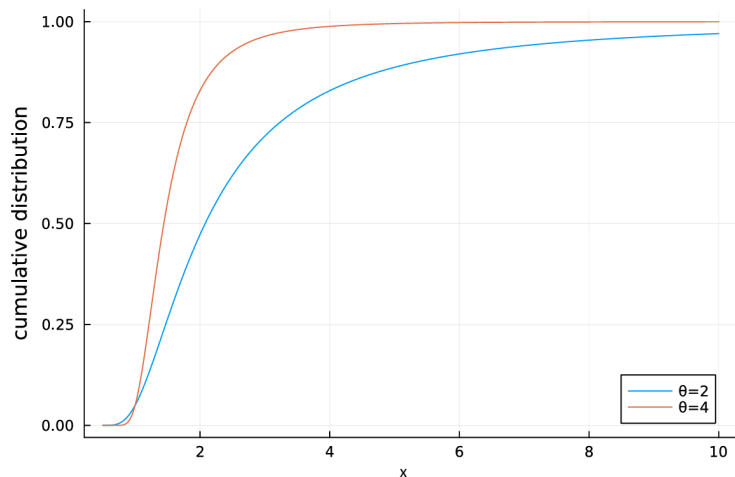
- ▶ Generally, the distribution function of the Fréchet distribution is

$$F(x) = e^{-\left(\frac{x-m}{s}\right)^{-\theta}},$$

where

- ▶  $s$  is the scale parameter,
  - ▶  $m$  is the location parameter.
- ▶ We have set  $s = 1$  and  $m = 0$ .

# Illustrations of the Fréchet distribution



The lower  $\theta$  is, the larger the variability is.

## Going back to the problem

- ▶ Because there are  $n$  locations, the infinitesimal individual observes  $n$  deterministic utilities

$$v_1, v_2, \dots, v_n,$$

and  $n$  preference shocks

$$\epsilon_1, \epsilon_2, \dots, \epsilon_n.$$

- ▶ She chooses her best location by solving

$$\max_{i \in \{1, \dots, n\}} v_i \epsilon_i. \tag{1}$$

## Probability that an individual lives in location $i$

- ▶ The probability that the infinitesimal individual chooses location  $i$  is

$$P[v_i \epsilon_i > v_j \epsilon_j \text{ for any } j \neq i].$$

- ▶ By integrating the Fréchet distribution, this probability is computed as

$$\frac{v_i^\theta}{\sum_{j=1}^n v_j^\theta}.$$

- ▶ See `frechet_choice_probability.pdf` for derivations.
- ▶ An advantage of extreme value distributions is that we can write down analytical expression of choice probabilities.

## Probabilities and shares

- ▶ So the probability that any infinitesimal individual lives in location  $i$  is

$$P[v_i \epsilon_i > v_j \epsilon_j \text{ for any } j \neq i] = \frac{v_i^\theta}{\sum_{j=1}^n v_j^\theta}.$$

- ▶ There are a unit mass of individuals.
- ▶ These individuals receive i.i.d. preference shocks.
- ▶ Therefore, the share of individuals in location  $i$  is also  $\frac{v_i^\theta}{\sum_{j=1}^n v_j^\theta}$ .
- ▶ Since the total population is 1, the population in location  $i$  is also  $\frac{v_i^\theta}{\sum_{j=1}^n v_j^\theta}$ .
- ▶ If the total population is  $L > 0$ , the population in location  $i$  is  $\frac{v_i^\theta}{\sum_{j=1}^n v_j^\theta} L$ .

# Welfare (1)

## Ex ante expected utility

- ▶ Computing ex post utility does not make a lot of sense.
- ▶ "Even though  $v_i$  is very low, I happened to draw a high  $\epsilon_i$ , so I very happily live in location  $i$ ."
- ▶ Rather, we consider ex ante expected utility as a measure of welfare.
- ▶ Before an individual draws preference shocks, she computes her expected welfare.
- ▶ But she knows she will live in a location that provides the highest utility (including the stochastic part).
- ▶ Then the ex ante expected utility is

$$E \left[ \max_{i \in \{1, \dots, n\}} v_i \epsilon_i \right].$$

## Welfare (2)

### Ex ante expected utility

- ▶ The ex ante expected utility is computed as

$$E \left[ \max_{i \in \{1, \dots, n\}} v_i \epsilon_i \right] = \Gamma \left( 1 - \frac{1}{\theta} \right) \left( \sum_{i=1}^n v_i^\theta \right)^{1/\theta},$$

where  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$  is the gamma function.

- ▶ See `frechet_expected_max.pdf` for derivations.
- ▶ This analytical expression for the "expectation of a maximum" (not the "maximum of expectations") is also an advantage of extreme value distributions.
- ▶ Since  $\Gamma \left( 1 - \frac{1}{\theta} \right)$  is constant, the ex ante expected utility is basically the  $p$ -norm of the vector of utilities in destinations.
  - ▶ For  $p > 1$ , the  $p$ -norm of a vector  $x = (x_1, \dots, x_n)$  is

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

## Gumbel (1)

- ▶ We have discussed the case where preference shocks follow the Fréchet distribution.
- ▶ But the title of today's lecture is "discrete choices and extreme value distributions."
- ▶ Another name of the Fréchet distribution is the type-2 extreme value distribution.
- ▶ There must be type-1.
- ▶ The type-1 extreme value distribution is referred to as the Gumbel distribution.
- ▶ The distribution function of the Gumbel distribution with location parameter  $\mu$  and scale parameter  $\sigma$  is

$$H(x) = e^{-e^{-(x-\mu)/\sigma}}. \quad (2)$$

- ▶ We can assume Gumbel preference shocks instead of Fréchet preference shocks.

## Gumbel (2)

- ▶ For Gumbel preference shocks, we assume additive preference shocks instead of multiplicative preference shocks.
- ▶ That is, individuals choose their best location solving

$$\max_i v_i + \epsilon_i,$$

where

- ▶  $v_i$  is, as before, the deterministic part of the utility in  $i$ ,
  - ▶  $\epsilon_i$  is the stochastic part that follows (2) independently across locations and individuals.
- ▶ Or, oftentimes spatial economists normalize  $\mu = 0$  and  $\sigma = 1$  in (2) and assume that individuals solve

$$\max_i v_i + \nu \epsilon_i. \tag{3}$$

- ▶ See Caliendo et al. (2019) for a full general equilibrium model including this problem.
- ▶ Their model is dynamic, though.

## Gumbel (3)

- ▶ Here we consider aggregate consequences of individual problem (3).
- ▶ Then the population share of location  $i$  is

$$\frac{\exp\left(v_i^{1/\nu}\right)}{\sum_j \exp\left(v_j^{1/\nu}\right)}.$$

- ▶ And the ex ante expected utility is

$$\nu \log \left( \sum_i \exp\left(v_i^{1/\nu}\right) \right).$$

# Fréchet vs Gumbel

- ▶ Compare the maximization with the Fréchet distribution (1) and the one with the Gumbel distribution (3).
- ▶ Fréchet: preference shocks are multiplicative.
- ▶ Gumbel: preference shocks are additive.
- ▶ These assumptions lead to handy expressions for shares and welfare.

## References I

- Anderson, S. P., De Palma, A., and Thisse, J.-F. (1987). The CES is a discrete choice model? *Economics Letters*, 24(2):139–140.
- Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor market power. *American Economic Review*, 112(4):1147–93.
- Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the China trade shock. *Econometrica*, 87(3):741–835.
- Verboven, F. (1996). The nested logit model and representative consumer theory. *Economics Letters*, 50(1):57–63.