Notes on *Discrete Choice Methods with Simulation* (Train, 2009)

Chapter 1. Introduction: choice probabilities and integration

- An agent (i.e., person, firm, decision maker) faces a choice, or a series of choices over time, among a set of options.
- Denote the outcome of the decision(s) in any given situation as y, indicating the chosen option or sequence of options.
- Our goal is to understand the behavioral process that leads to the agent's choice. There are factors that collectively determine the agent's choice: observable factors x and unobservable factors ε. These factors relate to the agent's choice through a function $y = h(x, \varepsilon)$, which is called the behavioral process. Given x and ε, the choice of the agent is fully determined.
- However, what researchers could observe are observable factors x and the probability of any particular outcome derived. Since ε is not observed, the agent's choice is actually not deterministic for researchers.
- Considering the unobservable factors with density $f(\varepsilon)$, the probability that the agent chooses a particular outcome from the set of all possible outcomes is

  $$P\left(y \mid x\right) = \Pr\left\{\varepsilon \quad \text{s.t.} \quad h(x, \varepsilon) = y\right\},$$

  or we can define a indicator function $\mathbb{I}[h(x, \varepsilon) = y]$ to express the probability in a more usable form:

  $$P\left(y \mid x\right) = \int \mathbb{I}\left[h\left(x, \varepsilon\right) = y\right] f\left(\varepsilon\right) \mathrm{d}\varepsilon.$$

- To calculate this probability, the integral must be evaluated, which leads to three types of probabilities:
  - *Complete closed-form expression:* Multinomial logit, nested logit, and ordered logit.
  - *Complete simulation:* Probit.
  - *Partial simulation, partial closed form:* Mixed logit.