1 Estimating the preference for safety features in cars

Tool: Multinomial logit and mixed logit models.

The Analytics Edge: Companies such as General Motors carry out conjoint studies with customers to understand the tradeoff (valuation) of different attributes that make up a product or a service. Using data or preferences for safety features in new vehicles and then building models of discrete choice, the company can obtain estimates on the valuation of attributes. This provides important information that can be used to infer the effect of introducing new products in the new market, pricing the product and designing the product features.

The models incorporate the comparison of attributes across alternatives and can also be used to capture heterogeneity in the choice modelling process.

1.1 Mixed logit model (or, random parameters logit model)

A Mixed logit model is often used to capture additional features that multinomial logit cannot capture such as random taste variation.

- Standard logit: $U_{ik} = \boldsymbol{\beta}^\mathsf{T} \boldsymbol{x}_{ik} + \epsilon_{ik}$.
- Mixed logit: $\tilde{\boldsymbol{\beta}}$ is modeled as a random parameter, $U_{ik} = \tilde{\boldsymbol{\beta}}^\mathsf{T} \boldsymbol{x}_{ik} + \epsilon_{ik}$.

Standard multinomial logit	Mixed logit
$\Pr(Y_i = k) = \dfrac{e^{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_{ik}}}{\sum_{l=1}^{K} e^{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_{il}}}$	$\Pr(Y_i = k) = \displaystyle\int \dfrac{e^{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_{ik}}}{\sum_{l=1}^{K} e^{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}_{il}}} f(\boldsymbol{\beta})\, \mathrm{d}\boldsymbol{\beta}$

For the mixed logit model, $f(\boldsymbol{\beta})$ is the probability density function of $\tilde{\boldsymbol{\beta}}$. This leads to an integral over logit probabilities.

- Mixed logit is computationally more challenging to solve due to the use of simulation optimization methods.
- When we write down the log-likelihood equation and try to maximize it with respect to the parameter choice, the problem is no longer a concave maximization problem. In this setting, finding a global optimum might not be easy.
- From an estimation perspective, the goal is to find the parameters $\boldsymbol{\theta}$ that define the density function $f(\boldsymbol{\beta}|\boldsymbol{\theta})$ where the functional form $f(\cdot)$ is given but parameters $\boldsymbol{\theta}$ are unknown.

- The simulation estimation technique works as follows:

  1. Make an initial hypothesis about the parameter $\boldsymbol{\theta}$.
  2. Draw R samples from this distribution $f(\cdot|\boldsymbol{\theta})$, call them $\boldsymbol{\beta}_r, r = 1, \ldots, R$ and compute

$$P_{ik}^r = \frac{e^{\boldsymbol{\beta}_r^\top \boldsymbol{x}_{ik}}}{\sum\limits_{l=1}^{K} e^{\boldsymbol{\beta}_r^\top \boldsymbol{x}_{il}}}.$$

  3. Now we approximate $\Pr(Y_i = k)$ by the average:

$$\Pr(Y_i = k) \approx \overline{P}_{ik} = \frac{1}{R} \sum_{r=1}^{R} P_{ik}^r.$$

  4. Compute the (simulated) log-likelihood for the probabilities:

$$SLL = \sum_{i=1}^{n} \sum_{l=1}^{K} z_{ik} \log \overline{P}_{ik}$$

     where $z_{ik} = 1$ if ith individual chose the kth item.
  5. Repeat the process with a new choice of $\boldsymbol{\theta}$ and maximize (optimize).

- For mixed logit with repeated choices (panel data), where i: individual, k: alternative, t: observation

$$\Pr(Y_{i1} = k_1, \ldots, Y_{iT} = k_T) = \int \prod_{t=1}^{T} \left(\frac{e^{\boldsymbol{\beta}^\top \boldsymbol{x}_{ik_t t}}}{\sum_{l=1}^{K} e^{\boldsymbol{\beta}^\top \boldsymbol{x}_{ilt}}} \right) f(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

- In panel data, we need to account for the fact that the errors are correlated for the same individual over time.

1.2 Willingness to pay

In discrete choice modeling, where utilities of items are modeled with different attributes (including price), the effects of the variables may be gauged by a quantity called the *willingness to pay.*

Suppose β_1 is the coefficient for attribute 1 and β_2 is the coefficient of attribute 2 (Price). Note that β_2 will be typically negative since more the price, lesser the utility. Say β_1 is positive. For a particular level of attribute 1, say, x_1 and a price x_2, we have utility

$$U = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \text{other terms}.$$

The question is that for a one unit increase in attribute 1, how much price are we willing to pay (wtp) so that the utility remains the same. Suppose this amount is Δ. Then

$$U = \beta_0 + \beta_1(x_1 + 1) + \beta_2(x_2 + \Delta) + \text{other terms}$$

Hence we obtain, the *willingness to pay* for unit increase in attribute 1 is

$$wtp = \Delta = -\frac{\beta_1}{\beta_2}.$$