Likert Scale Dual Response in Conjoint Analysis

Prachi Bhalerao Dan Yavorsky Geoffery Zheng

2024-07-02

Abstract

Write the abstract here …

1 Introduction

Lorem ipsum

2 Model

2.1 General Specification

Consumer $i = 1, \dots, N$ derives utility from good $j \in \mathcal{J}_i = \{0, \dots, J_i\}$ with utility u_{ij} given by

$$u_{ij} = h(\boldsymbol{x}_j, \boldsymbol{\beta}_i) + \eta_{ij}.$$

where $\boldsymbol{x}_j$ is a vector of good characteristics, $\boldsymbol{\beta}_i$ is a vector of consumer-specific taste parameters, and η_{ij} encapsulates factors known to the consumer but not to the researcher that affect the consumer's utility and are modeled as $\eta_{ij} \sim \text{Gumbel}\,(0, 1)$ with η_{ij} assumed to be indepedent of both x_j and $\eta_{ij'}$ for $j' \neq j$. We take $h(\boldsymbol{x}_j, \boldsymbol{\beta}_i) = \boldsymbol{x}_j'\boldsymbol{\beta}_i$ but this specification is not required.

The "zero"-th (or "outside") good is special and is associated with a zero vector of good characteristics ($\boldsymbol{x}_0 = \boldsymbol{0}$) such that $h(\boldsymbol{x}_0, \boldsymbol{\beta}_i) = 0$. Consumers observe $\boldsymbol{x}_j$ and η_{ij} for all "inside" goods ($j > 0; j \in \mathcal{J}_i$), but they *do not* observe η_{i0}.[1]

Consumers first report their preferred inside good among, which is given by

$$j_i^* = \arg \max_{j>0; j \in \mathcal{J}_i} u_{ij}.$$

Let u_i^* denote the utility of good j_i^* for consumer i and let $t_i = e_{j_i^*}$ indicate the "one-hot" encoding of the most-preferred inside good (ie, t_i is a vector with $J_i - 1$ zeros and 1 one).

[1]This can be motivated by a framework in which $u_0 = 0$ and $u_{ij} = x_j{}'\beta_i + \eta_{ij} - \eta_0$ for $j > 0; j \in \mathcal{J}_i$. Here, η_0 captures the consumer's uncertainty about their future tastes. Given that utilities are ordinal and η_0 is a common shock, it plays no role in the choice among the most preferred inside good $j^* \in \mathcal{J}_i$.]

Second, consumers report a value y_i on a discrete qualitative scale $w \in \mathcal{W} = \{1, \dots, W\}$ to indicate the probability that they prefer good j_i^* to the outside good 0. Consumers know u_i^*, but do not know η_{i0}, and thus this probability is given by

$$p_i = \Pr\left(\eta_{i0} < u_i^*\right).$$

Each consumer reports the interval $y_i = w$ into which p_i falls

$$y_i = w \text{ s.t. } p_i \in \left[\alpha_{w-1}, \alpha_w\right).$$

As is well known,[2] u_i follows a Gumbel distribution with location parameter $\overline{\mu}_i$ and scale parameter 1, where

$$\overline{\mu}_i = \ln\left(\sum_{j \in \mathcal{J}} \exp\left(X_j{}' \beta_i\right)\right). \tag{1}$$

Under these assumptions, we have $p_i = F\left(u_i^*\right)$ with F as the $\text{Gumbel}\left(0, 1\right)$ distribution.

The consumer's report of $y_i = w$ is therefore equivalent to reporting that

$$u_i^* \in \left[F^{-1}_{\text{Gumbel}(0)}\left(\alpha_{w(i)-1}\right), F^{-1}_{\text{Gumbel}(0)}\left(\alpha_{w(i)}\right)\right),$$

where we omit the common scale parameter for brevity.

This occurs with probability

$$\begin{aligned}
\Pr\left(u_i^* \in \mathcal{W}_{w(i)}\right) &= F_{\text{Gumbel}(\overline{\mu})}\left(F^{-1}_{\text{Gumbel}(0)}\left(\alpha_{w(i)}\right)\right) - F_{\text{Gumbel}(\overline{\mu})}\left(F^{-1}_{\text{Gumbel}(0)}\left(\alpha_{w(i)-1}\right)\right) \\
&= \left(\alpha_{w(i)}\right)^{\exp(\overline{\mu})} - \left(\alpha_{w(i)-1}\right)^{\exp(\overline{\mu})}.
\end{aligned}$$

This is $p\left(w_i \mid \alpha, \beta\right)$, the conditional likelihood of w_i.[3]

2.2 Comment: Relation to the (Brazell et al. 2006) Dual Response Model

Another common framework when soliciting consumer preferences is to directly ask consumers if $V_i \geq 0$. This implicitly assumes that $\eta_{i0} = 0$, so that consumers deterministically know whether or not the ''inside'' good j_i^* is preferred to the outside good. As we have already noted, our framework can nest that standard case by assuming that, rather than following a standard Gumbel distribution, η_0 is instead a degenerate distribution. In such a case, all but 2 of the W partitions are empty, and the non-empty partitions are $\mathcal{W}_0 = \{0\}$ and $\mathcal{W}_W = \{1\}$, as $p_i \in \{0, 1\}$.

We acknowledge that this requires a slight abuse of notation, as $\mathcal{W}_w$ was defined above using left-closed intervals. These have the advantage of being invertible under the inverse-CDF mapping. In the degenerate case, we instead have $p\left(w_i = W \mid \alpha, \beta\right) = \mathbb{P}\left(\eta_{i0} < V_i\right) = \mathbb{P}\left(V_i > 0\right) = 1 - \exp\left(-\exp\left(\overline{\mu}\right)\right)$.

[2] For example, see (McFadden 1981) and (Cardell 1997).

[3] While the parametrization of $\eta_0 \sim \text{Gumbel}\left(0, 1\right)$ preserves symmetry among the $J+1$ goods and is thus a natural choice, the framework can easily accommodate an alternative distribution for η_0. For example, one could use an affine function of individual characteristics to accommodate individual-level variation in the propensity to prefer the outside good.

2.3 Comment: Nesting a Single-Good Specification

Suppose there are two goods: good j and good 0. For example, suppose good j is Diet Coke and good 0 is the "outside option" of not purchasing the beverage.

These goods provide utility of

- $U_j = x_j'\beta + \eta_j$ and
- $U_0 = x_0'\beta + \eta_0$.

x is observed by both the consumer and the researcher (eg, the price x_j of the Diet Coke and the price of not making a purchase $x_0 = 0$); β are taste parameters known to the consumer, but not to the researcher that are to be estimated (eg, her price sensitivity).

η encapsulates factors known to the consumer but not to the researcher that affect the consumer's utility (eg, the positive or negative "status" from being observed purchasing or consuming the Diet Coke, or from not purchasing and consuming a beverage). From the researcher's perspective, η_j and η_0 are assumed to be independent of x and modeled as random variables with cumulative distribution functions $F(\eta_j)$ and $F(\eta_0)$.

We define the difference in utility as $U^* = U_j - U_0 = x_j'\beta + \eta^*$ where $x_0 = 0$ and $\eta^* = \eta_j - \eta_0$.

The consumer does not report U^* but rather reports y, a censoring of U^* into one of W discrete qualitative scale values $w = 1, 2, \dots, W$. Suppose, for example, that the consumer reports the middle level ($y = 2$) out of three available (labeled, "unlikely" for $w = 1$, "somewhat likely" for $w = 2$, and "very likely" for $w = 3$).

The W levels of the qualitative scale are separated at values μ_w such that $-\infty < \mu_1, \mu_2, \dots, \mu_W = \infty$, with $\mu_w, w = 1, 2, \dots, W$ as parameters to be estimated.

Then we have that

$$\begin{aligned}
\Pr\left(y = w\right) &= \Pr\left(\mu_{w-1} < U^* < \mu_w\right) \\
&= \Pr\left(\mu_{w-1} < x_j'\beta + \eta^* < \mu_w\right) \\
&= \Pr\left(\mu_{w-1} - x_j'\beta < \eta^* < \mu_w - x_j'\beta\right) \\
&= F(\mu_w - x_j'\beta) - F(\mu_{w-1} - x_j'\beta)
\end{aligned}$$

Define $t_w \in \{0, 1\}$ as indicators with $t_w = 1$ when $y = w$. The individual likelihood is then

$$\prod_{w=1}^{W} = \left[F(\mu_w - x_j'\beta) - F(\mu_{w-1} - x_j'\beta)\right]^{t_w}$$

2.4 Sketch of Estimation

- Specify hyper-parameters governing the prior distribution of $(\boldsymbol{\pi}, \boldsymbol{\beta})$. There is a relatively large amount of flexibility in the prior distribution over β. $\boldsymbol{\pi} \sim$ Dirichlet deterministically maps to $\boldsymbol{\alpha}$.

- (First Branch) Given draws of (α, β), the probability that consumer i reports j_i^* is given by a softmax

$$p\left(j_i^* \mid \alpha, \beta\right) = p\left(j_i^* \mid \beta\right) = \frac{\exp\left({X_j}'\beta_i\right)}{\sum_{j' \in \mathcal{J}} \exp\left({X_{j'}}'\beta_i\right)}$$

- (Second Branch) Given (α, β), the probability that consumer i reports w_i is given by

$$p\left(w_i \mid \alpha, \beta, j_i^*\right) = p\left(w_i \mid \alpha, \beta\right) = \left(\alpha_{w(i)}\right)^{\exp(\overline{\mu})} - \left(\alpha_{w(i)-1}\right)^{\exp(\overline{\mu})}$$

  where $\overline{\mu}_i\left(\beta_i\right)$ is a function of consumer i's tastes β_i and the design matrix X (@eq-max-mu). Note that the observed choice j_i^* is irrelevant for this likelihood.

Thus the overall likelihood (conditional on some draw of parameters) is

$$p\left(\left(j_i^*, w_i\right) \mid \alpha, \beta\right) = p\left(w_i \mid \alpha, \overline{\mu}_i\left(\beta\right)\right) \times p\left(j^* \mid \beta\right).$$

Assume some universal partitioning of the unit interval into W disjoint intervals with cutoffs denoted by the $(W+1)$-dimensional vector $\boldsymbol{\pi} \in \Delta^{W-1}$. Let $\boldsymbol{\alpha}$ denote the partial sums of $\boldsymbol{\pi}$, so that

$$\boldsymbol{\alpha} = \left\{0, \quad \pi_1, \quad \pi_1 + \pi_2, \quad \ldots, \quad \underbrace{\textstyle\sum_{i=1}^{w} \pi_i}_{\alpha_w}, \quad \ldots, \quad 1\right\}$$

$$\mathcal{W}_w = \left[\alpha_{w-1}, \alpha_w\right)$$

$$\bigcup_{w=1}^{W} \mathcal{W}_i = [0, 1)$$

3 Simulation Study

add

4 Empirical Analysis

add

5 Discussion

Differs from (Brazell et al. 2006)

6 Conclusion

add

References

Brazell, Jeff D., Christopher G. Diener, Ekaterina Karniouchina, William L. Moore, Válerie Séverin, and Pierre-Francois Uldry. 2006. "The No-Choice Option and Dual Response Choice Designs." *Marketing Letters* 17 (4): 255–68. https://doi.org/10.1007/s11002-006-7943-8.

Cardell, N. Scott. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory* 13 (2): 185–213. https://doi.org/10.1017/s0266466600005727.

McFadden, Daniel L. 1981. "Structural Discrete Probability Models Derived from Theories of Choice." In, 198–272. The MIT Press.