Profile Ranking Adaptive Choice-Based Conjoint Analysis: A Complementary Approach to Utility-Based Analysis of Small Populations

Abstract

To analyze Adaptive Choice-Based Conjoint (ACBC) survey samples from small populations, a new methodology called profile ranking based ACBC (PR-ACBC) is proposed as a complement to utility based ACBC. PR-ACBC offers a form of validation especially useful for small survey data with high variances in partworth utilities. Without requiring knowledge of partworth utilities, PR-ACBC deduces from sample data the maximum likelihood population ranking for known population sizes using a multivariate hypergeometric distribution, and for unknown population sizes using the Lagrange multiplier based optimization. A population ranking interval can easily be computed for each profile as an interval in which the population profile ranking must lie. Various distance measures from statistical ranking theory are used to analyze profile decomposition (importance of attribute levels in choice tasks), and multidimensional scaling (MDS) for visual representation of attribute importances of particular value in comparing sample sub-groups. The methodology of PR-ACBC is introduced using a toy survey, and its application by a recent survey administered to faith-based and non-faith based disaster relief organizations belonging to the National Voluntary Organizations Active in Disaster (VOAD).

1 Introduction

Adaptive Choice-Based Conjoint (ACBC) analysis surveys are a widely-utilized, well-developed and highly effective type of conjoint analysis (Orme and Chrzan, 2017). While the Max-Diff approach to select the best and worst among several profiles has generated a great deal of research interest, we focus on ACBC surveys whose choice tasks are designed with the simplest choice between just two concepts. Following Sawtooth's Lighthouse survey creation tool, we structure our choice-task stage as a tournament beginning with 16 profiles close to the respondent's #1 profile called "Build Your Own" (BYO). For large samples, use of a sophisticated statistical method such as hierarchical-Bayesian Markov-Chain Monte-Carlo (HB MCMC) simulation (Rossi et. al. 2005) is extremely effective to estimate partworth utilities and their variances. In the case of very small samples (eg. $n \leq 15$) from a small population (eg. $N \leq 50$), the variances in partworth utility may hamper both accurate prediction of choice experiments and ranking of attribute importances. Profile ranking (PR-) ACBC is therefore introduced as a method serving as a validity check for profile choice prediction based on partworth levels, as well as attribute importance rankings derivable from the partworth utilities. PR-ACBC utilizes distance measures from statistical ranking theory for attribute decomposition as an alternative to partworth utilities for choice predictions, and multi-dimensional scaling (MDS) as a means to assess attribute importances (Alvo and Yu 2014).

In Section 2, we introduce basic PR-ACBC methodology by means of a generic survey with only 4 profiles constructed from 2 attributes each with 2 levels. We begin with a fundamental observation that the exact sample profile rankings directly obtainable from choice tournament data can not be predicted by multiple linear regression of part-worth utilities. PR-ACBC then proceeds to analyze survey tournament data without requiring any knowledge of partworth utilities. Maximum likelihood estimate (MLE) population rankings for known population sizes are obtainable by discrete multivariate hypergeometric distribution (Oberhofer and Kaufman 1987), and for unknown population sizes by multivariable calculus optimization using Lagrange multipliers (Stewart 2016). Population ranking intervals (PRIs) are easily computed from sample tournament data as one-dimensional intervals which must contain the population profile rankings. Statistical ranking theory distance measures are then used for profile attribute decomposition (akin to partworth levels) and multi-dimensional scaling (MDS) for attribute importances. These are important for comparing and contrasting sample subgroups. In Section 3 we illustrate PR methodology using a recent ACBC survey deployed to both faith-based and non-faith based disaster relief organizations. The context motivating this methodological study is a sequel to a novel application of ACBC in disaster-response research (Gralla et. al. 2014).

2 PR-ACBC Methodology

2.1 Simple Example

Consider a generic ACBC survey with just 2 attributes each having 2 levels. We designate the 4 possible profiles $A = 11, B = 12, C = 21, D = 22$, where $X = x_1 x_2$ designates that profile X has level x_1 for the first attribute and level x_2 for the second attribute. We consider a population with $N = 8$ possible respondents from which sample data for $n = 4$ randomly selected respondents is obtained by a head-to head choice tournament of the 4 profiles. A sample tournament outcome is shown in Figure 1

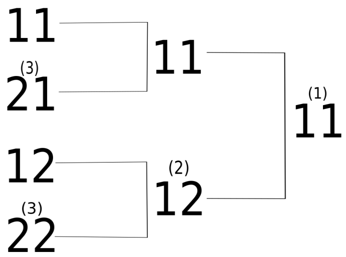

Figure 1: One possible tournament outcome ranks profile A=11 first, B=12 second, and the remaining two profiles C=21 and D=22 third.

For this respondent, profile A=11 is ranked 1 (tournament winner), profile B=12 is ranked 2 (runner-up), and profiles C=21 and D=22 are both ranked 3. Note that each of the $n = 4$ sample respondents will have a tournament outcome, which we can compile into an *Individual Ranking Table (IRT)*, an example of which is shown in Table 1.

	Rank		
Respondent	1	2	3
1	A	B	C,D
2	A	C	B,D
3	B	C	A,D
4	A	C	B,D

Table 1: Individual Ranking Table. Tournament outcomes for a sample of 4 respondents.

From this data, it is easy to compute the *Sample Ranking Table (SRT)* which indicates how many respondents assigned a particular rank to each profile (Table 2).

Given any profile X and a sample of size $n = 4$, the *sample profile ranking (SPR)* $\rho_n(X) = \rho_4(X)$ is obtained as an arithmetic average of the $n = 4$ respondents. For example, for A=11, $\rho_4(A) = [3(1) + 1(3)]/4 = 1.5$, and for C=12 it is $\rho_4(C) = [3(2) + 1(3)]/4 = 2.25$. In this case, profile A=11 has the best sample profile ranking, profiles B=12 and C=21 are tied for second, and profile D=22 comes in fourth.

	Rank			
Profile	1	2	3	SPR
A	3	0	1	1.5
B	1	1	2	2.25
C	0	3	1	2.25
D	0	0	4	3

Table 2: Sample Ranking Table (SRT). The i, j entry of the 4x3 matrix forming the body of the table shows the number of respondents who assigned to the profile in the margin of row i, the rank shown by the header in column j. The sample profile ranking (SPR) is the arithmetic average of the sample respondent rankings.

Our main question is "*what can we infer from the SRT about the population ranking table (PRT)?*", the latter meaning the profile ranking table for all $N = 8$ population respondents. To begin, note that there are essentially 12 different possible tournament outcomes which are essentially determined by specifying the winner and runner-up (Table 3):

	Rank		
Outcome	1	2	3
O_1	A	B	C,D
O_2	A	C	B,D
O_3	A	D	B,C
O_4	B	A	C,D
O_5	B	C	B,D
O_6	B	D	A,C
O_7	C	A	B,D
O_8	C	B	A,D
O_9	C	D	A,B
O_{10}	D	A	B,C
O_{11}	D	B	C,D
O_{12}	D	C	A,B

Table 3: The 12 possible tournament outcomes.

Each member of the population with size $N = 8$ has an outcome belonging to the set $\mathcal{S} = \{O_1, ..., O_{12}\}$. In other words, the outcomes for a population with $N = 8$ consists of $N = 8$ draws (with replacement) from a hat containing the 12 outcomes in the set $\mathcal{S}$.

	Rank			
Respondent	1	2	3	Outcome
1	A	B	C,D	O_1
2	A	C	B,D	O_2
3	B	C	A,D	O_5
4	A	C	B,D	O_2
5	unknown			unknown
6	unknown			unknown
7	unknown			unknown
8	unknown			unknown

Table 4: The individual ranking table for the entire population must consist of the 4 sample outcomes plus an additional 4 outcomes which are unknown.

Returning to our example, as shown in Table 4, we note that *any* outcomes for respondents 5, 6, 7 and 8 could in principle have resulted in the observed sample outcomes for respondents 1, 2, 3 and 4. However, not all choices for the unknown outcomes would have the same probability of generating a sample with $n = 4$ such that two of the outcomes are O_2, and one each of the outcomes are O_1 and O_5. For example, let us consider first the case where the unknown population outcomes are exactly equal to the observed sample. That is, the outcomes for the unknown 4 population members also consist of two with outcome O_2 and one each with outcomes O_1 and O_5. As a result, the total population consists of four with outcome O_2 and two each with outcomes O_1 and O_5. The resulting *population ranking table* (PRT) shown in Table 5 is in this case obtained by simply doubling each entry in the body of the SRT (Table 2).

	Rank			
Profile	1	2	3	PPR
A	6	0	2	1.5
B	2	2	4	2.25
C	0	6	2	2.25
D	0	0	8	3

Table 5: Population Ranking Table (PRT) for a population $N = 2n = 8$ obtained by doubling each entry in the body of the SRT of Table 2. In this case, the population profile rankings (PPR) are identical to the sample profile rankings (SPR)

2.2 A Fundamental Observation

In this section we explain why least squares multiple regression (LSRM) will not in general give exact sample profile rankings. In other words, part-worth utilities can only approximate sample rankings.

2.3 Analytic Approach

Least squares multiple regression (LSRM) can be used to predict sample profile rankings based on individual sample outcomes as we will now explain using our simple 2 attribute, 2 level survey and generic sample data for 4 respondents. Table 6 gives the dataset $\{(U_i, V_i, Y_i)\}$ $(i = 1, ..., 16)$ where

$$\begin{aligned} U_i &= 1 \text{ if attribute 1 has level 1, and 0 if it has level 2} \\ V_i &= 1 \text{ if attribute 2 has level 1, and 0 if it has level 2} \\ Y_i &= \text{ Respondent's ranking of a profile with } U = U_i,\ V = V_i. \end{aligned}$$

This dataset has certain properties:

- Each column consists of a respondent's profile rankings, and hence contains the values $1, 2, 3, 3$ in any order.
- Table 6 can also be represented in the form of Table 7, by which we see that

$$\sum U_i = \sum V_i = \sum U_i^2 = \sum V_i^2 = 8,$$

		Output Ranking Data			
U	V	Respondent 1	Respondent 2	Respondent 3	Respondent 4
1	1	Y_1	Y_2	Y_3	Y_4
1	0	Y_5	Y_6	Y_7	Y_8
0	1	Y_9	Y_{10}	Y_{11}	Y_{12}
0	0	Y_{13}	Y_{14}	Y_{15}	Y_{16}

Table 6: Dataset ranking structure by respondents.

and

$$\sum U_i V_i = 4,$$

where the symbol $\sum$ represents $\sum_{n=1}^{16}$.

U	V	Rank
1	1	Y_1
1	1	Y_2
1	1	Y_3
1	1	Y_4
1	0	Y_5
1	0	Y_6
1	0	Y_7
1	0	Y_8
0	1	Y_9
0	1	Y_{10}
0	1	Y_{11}
0	1	Y_{12}
0	0	Y_{13}
0	0	Y_{14}
0	0	Y_{15}
0	0	Y_{16}

Table 7: Dataset's ranking structure with respondents combined.

Using least squares multiple linear regression (LSMR) on the dataset in Table 7, we estimate each sample profile ranking Y_i as $\hat{Y}_i$:

$$\hat{Y}_i = c_0 + c_1 U_i + c_2 V_i,$$

where the regression coefficients c_0, c_1, c_2 are determined by minimizing the sum of squared residuals (SSR):

$$SSR = \sum_{n=1}^{16} (Y_i - \hat{Y}_i)^2 = \sum_{n=1}^{16} (Y_i - (c_0 + c_1 U_i + c_2 V_i))^2.$$

To minimize the SSR, we set the partial derivatives with respect to c_0, c_1 and c_2, equal to zero:

$$\frac{\partial SSR}{\partial c_0} = \frac{\partial SSR}{\partial c_1} = \frac{\partial SSR}{\partial c_2} = 0.$$

This yields the linear system:

$$\begin{cases} nc_0 + c_1 \sum U_i + c_2 \sum V_i = \sum Y_i \\ c_0 \sum U_i + c_1 \sum U_i^2 + c_2 \sum U_i V_i = \sum U_i Y_i \\ c_0 \sum V_i + c_1 \sum U_i V_i + c_2 \sum V_i^2 = \sum V_i Y_i \end{cases},$$

which is equivalent to the matrix equation:

$$\begin{bmatrix} n & \sum U_i & \sum V_i \\ \sum U_i & \sum U_i^2 & \sum U_i V_i \\ \sum V_i & \sum U_i V_i & \sum V_i^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum U_i Y_i \\ \sum V_i Y_i \end{bmatrix}.$$

Simplifying the sums and using Cramer's Rule we obtain the regression coefficients c_0, c_1 and c_2:

$$c_0 = \frac{\begin{bmatrix} \sum Y_i & 8 & 8 \\ \sum U_i Y_i & 8 & 4 \\ \sum V_i Y_i & 4 & 8 \end{bmatrix}}{256} = \frac{\begin{bmatrix} \sum Y_i & 2 & 2 \\ \sum U_i Y_i & 2 & 1 \\ \sum V_i Y_i & 1 & 2 \end{bmatrix}}{16},$$

$$c_1 = \frac{\begin{bmatrix} 16 & \sum Y_i & 8 \\ 8 & \sum U_i Y_i & 4 \\ 8 & \sum V_i Y_i & 8 \end{bmatrix}}{256} = \frac{\begin{bmatrix} 4 & \sum Y_i & 2 \\ 2 & \sum U_i Y_i & 1 \\ 2 & \sum V_i Y_i & 2 \end{bmatrix}}{16}, \text{ and}$$

$$c_2 = \frac{\begin{bmatrix} 16 & 8 & \sum Y_i \\ 8 & 8 & \sum U_i Y_i \\ 8 & 4 & \sum V_i Y_i \end{bmatrix}}{256} = \frac{\begin{bmatrix} 4 & 2 & \sum Y_i \\ 2 & 2 & \sum U_i Y_i \\ 2 & 1 & \sum V_i Y_i \end{bmatrix}}{16}.$$

The LSMR predicted profile rankings $\hat{Y}_{uv}$ are given by:

$$\hat{Y}_{11} = c_0 + c_1 + c_2 = \frac{2\sum U_i Y_i + 2\sum V_i Y_i - \sum Y_i}{16},$$

$$\hat{Y}_{10} = c_0 + c_1 = \frac{-6\sum V_i Y_i + \sum U_i Y_i + \sum Y_i}{16},$$

$$\hat{Y}_{01} = c_0 + c_2 = \frac{-2\sum U_i Y_i + 2\sum V_i Y_i + \sum Y_i}{16}, \text{ and}$$

$$\hat{Y}_{00} = c_0 = \frac{-2\sum U_i Y_i - 2\sum V_i Y_i + 3\sum Y_i}{16}.$$

The corresponding actual sample profile rankings obtained by averaging the respondent rankings are:

$$\bar{Y}_{11} = \frac{Y_1 + Y_2 + Y_3 + Y_4}{4},$$

$$\bar{Y}_{10} = \frac{Y_5 + Y_6 + Y_7 + Y_8}{4},$$

$$\bar{Y}_{01} = \frac{Y_9 + Y_{10} + Y_{11} + Y_{12}}{4}, \text{ and}$$

$$\bar{Y}_{00} = \frac{Y_{13} + Y_{14} + Y_{15} + Y_{16}}{4}.$$

We thus have the following theorem: *The sample profiles' LSMR predicted and actual rankings are the same if and only if the following four equations hold*

$$\text{(profile 11)}: 2\sum U_iY_i + 2\sum V_iY_i - \sum Y_i = 4(Y_1+Y_2+Y_3+Y_4), \quad (1)$$

$$\text{(profile 10)}: -6\sum V_iY_i + \sum U_iY_i + \sum Y_i = 4(Y_5+Y_6+Y_7+Y_8), \quad (2)$$

$$\text{(profile 01)}: -2\sum U_iY_i + 2\sum V_iY_i + \sum Y_i = 4(Y_9+Y_{10}+Y_{11}+Y_{12}), \text{and} \quad (3)$$

$$\text{(profile 00)}: -2\sum U_iY_i - 2\sum V_iY_i + 3\sum Y_i = 4(Y_{13}+Y_{14}+Y_{15}+Y_{16}). \quad (4)$$

Furthermore, by adding these equations, we obtain a corollary: *The following is a necessary condition for the LSMR predicted sample profile rankings to equal the sample profile rankings:*

$$\sum U_iY_i = \sum V_iY_i. \quad (5)$$

Table 9 gives an example where the average profile rankings and part-worth profile rankings are equal

		Respondents						
U	V	R 1	R 2	R 3	R 4	Actual Sample Rank	Predicted Sample Rank=$c_0+c_1U_1+c_2U_2$	Residual Error
1	1	1	1	3	1	1.5	3-.75(1)-.75(1)=1.5	0
1	0	2	3	1	3	2.25	3-.75(1)-.75(0)=2.25	0
0	1	3	2	2	2	2.25	3-.75(0)-.75(1)=2.25	0
0	0	3	3	3	3	3	3-.75(0)-.75(0)=3	0

Table 8: Predicted rankings for the sample outcomes in Table 1.

In this case, the regression coefficients are $c_0 = 3$, $c_1 = -.75$, $c_2 = -.75$. The actual sample profile rankings obtained by averaging the respondent rankings are equal to the predicted profile rankings obtained by LSMR, shown in Table 8.

Moreover, the equality (5) in the corollary holds:

$$\sum U_iY_i = \sum V_iY_i = 15.$$

The dataset in Table 10, shows 4 respondents whose profiles' predicted and actual rankings are not equal. In this case, the regression coefficients are $c_0 = 3$, $c_1 = -1$, $c_2 = -.5$. The actual profile rankings obtained by averaging the respondent rankings are not equal to the estimated profile rankings obtained by LSMR since all of the profile residuals are non-zero. This must be so since the corollary's condition (5) does not hold.

		Respondents						
U	V	R 1	R 2	R 3	R 4	Actual Rank	Predicted Rank=$c_0 + c_1U_1 + c_2U_2$	|Residual Error|
1	1	1	1	1	2	1.25	3-1(1)-.5(1)=1.5	0.25
1	0	2	3	3	1	2.25	3-1(1)-.5(0)=2	0.25
0	1	3	3	2	3	2.75	3-1(0)-.5(1)=2.5	0.25
0	0	3	2	3	3	2.75	3-1(0)-.5(0)=3	0.25

Table 9: LSMR predicted profile rankings for a sample will in general involve residual errors.

2.4 Geometric Interpretation

The conditions (12)-(15) for whether or not the LSMR profile ranking predictions are error-free may be understood geometrically by considering points in an $x_1x_2x_3$ coordinate system in which the x_1x_2 coordinates represent the profile and the x_3 coordinate the ranking. For our simple generic survey, if the four points representing the sample profile rankings are co-planar, there is no error; otherwise, the LSMR predicted profile rankings will have a residual error (Figure 2).

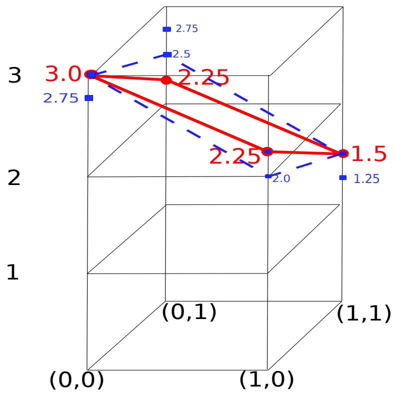

Figure 2: In special cases such as the one described in Table 9, the 4 points representing the sample profile rankings are co-planar ($Y_i = 3 - .75U_i - .75V_i$, shown in solid outline). Otherwise, as in Table 10, the utility-based sample profile rankings will have residual errors as shown by the sample rankings and plane $Y_i = 3 - U_i - .5V_i$ (dashed outline).

Such a geometric interpretation is not possible for surveys involving more than 2 attributes, in which case standard LSMR residual analysis indicates the error in sample profile rankings using regression coefficients.

2.5 Maximum Likelihood Estimation

2.5.1 Known Population Size

2.5.2 Unknown Population Size

Let us assume that each member of the population with $N = 8$ is equally likely to occur in our sample with $n = 4$. Assume further that p_i is the probability that a respondent's outcome is O_i. The probability p that the observed sample consisting of one O_1, two O_2's and one O_5 is

$$p = f(p_1, p_2, p_5) = \frac{4!}{1!2!1!}[p_1 p_2^2 p_5], \tag{6}$$

where $g(p_1, p_2, p_5) = p_1 + p_2 + p_5 = 1$. The values of p_1^*, p_2^*, and p_5^* which maximize $H(p_1, p_2, p_5) = \ln(f(p_1, p_2, p_5))$ (and hence also maximizes $p = f(p_1, p_2, p_5)$) are obtained using Lagrange multipliers:

$$\nabla H(p_1^*, p_2^*, p_5^*) \;\; = \;\; \lambda \nabla g(p_1^*, p_2^*, p_5^*),$$

and therefore

$$\begin{aligned} \frac{1}{p_1^*} &= \lambda \\ \frac{2}{p_2^*} &= \lambda \\ \frac{1}{p_5^*} &= \lambda. \end{aligned}$$

(The scalar quantity λ is called a Lagrange multiplier.) Using $p_1^* + p_2^* + p_5^* = 1$ gives $\frac{1}{\lambda} + \frac{2}{\lambda} + \frac{1}{\lambda} = 1$ and so $\lambda = 4$. Hence, the values $p_1^* = \frac{1}{4}, p_2^* = \frac{1}{2}$, $p_5^* = \frac{1}{4}$ maximize the probability of the observed sample outcomes. Thus, the PRT in Table 5 can be interpreted as the expected rankings for a population of size 8 which maximizes the likelihood of the observed sample outcomes. The likelihood of the observed sample would be lower for any other expected population rankings such as a PRT corresponding to a population whose response outcomes consist of two O_1's, two O_5's, three O_2's and the outcome O_7 (which reverses the winner and runner-up in outcome O_2). The revised PRT is shown in Table 10.

	Rank			
Profile	1	2	3	PPR
A	5	1	2	1.625
B	2	2	4	2.25
C	1	5	2	2.15
D	0	0	8	3

Table 10: A revised population ranking table (PRT_1) with $N = 2n = 8$ obtained by replacing an outcome O_2 with outcome O_7 in the population with $N = 8$ represented by Table 5. Note that the population ranking of profile A increases by .125, while profile C's decreases by .1.

In general, let n_k be the number of sample outcomes O_k $(k = 1, 2, ..., K)$ and let p_k be the probability that a respondent in the population has outcome O_k $(k = 1, 2, ..., K)$. The likelihood function $f(p_1, p_2, ..., p_K)$ giving the probability of observing the sample values $n_1, ..., n_K$ is given by

$$f(p_1, ..., p_K) = \frac{n!}{n_1! n_2! \cdots n_K!} \prod_{k=1}^{K} p_k^{n_k}, \tag{7}$$

with $\sum_{k=1}^{K} n_k = n$ and $\sum_{k=1}^{K} p_k = 1$. We seek to find the values $p_1^*, ..., p_K^*$ which maximize the likelihood function f, or equivalently, the log-likelihood function

$$H(p_1, ..., p_K) = \ln f = \ln(n!) - \sum_{k=1}^{K} n_k! + \sum_{k=1}^{K} n_k \ln(p_k), \tag{8}$$

subject to the constraint $g(p_1, ..., p_K) = p_1 + p_2 + ... + p_K = 1$. Properties of gradients imply that the optimal values p_i^* must satisfy

$$\nabla H(p_1^*, ..., p_K^*) = \lambda \nabla g(p_1^*, ..., p_K^*). \tag{9}$$

It follows that for $k = 1, ..., K$,

$$\frac{n_k}{p_k^*} = \lambda. \tag{10}$$

Hence, $n = \sum_{k=1}^{K} n_k = \lambda \sum_{k=1}^{K} p_k^* = \lambda$, and so the probabilities $p_k^* = \frac{n_k}{n}$ give the maximum likelihood of the observed sample outcomes n_k $(k = 1, 2, ..., K)$. For any sample of size n and number n_k of observed outcomes O_k $(k = 1, 2, ...K)$, the maximum likelihood probabilities $p_k^* = \frac{n_k}{n}$ indicate that for a population of size N, the expected number N_k of outcomes O_k is given by $E(N_k) = p_k N$. A maximum-likelihood population could be simulated by augmenting the observed n sample outcomes, where the probability of outcome O_k at each draw is given by p_k. For a large number of such randomly constructed populations of size N, for each k the average number of population outcomes O_k is approximately $p_k N$.

2.6 Population Ranking Intervals

Maximum likelihood provides some insight into the expected population profile rankings. In this section, given ranking data for a sample consisting of n survey respondents selected at random from a population with $N > n$ respondents, we will show how to construct confidence intervals which are guaranteed to include the population profile rankings.

2.7 Population Ranking Range

Returning to our simple example, in which each profile is ranked 1, 2, or 3, let $\rho_k(X)$ denote the ranking of profile X based on tournament results for k respondents. Note that for any population of k respondents $(n \leq k \leq N)$ which contains an observed sample of size n, the following inequality must hold:

$$\frac{k + n(\rho_n(X) - 1)}{k} \leq \rho_k(X) \leq \frac{3k + n(\rho_n(X) - 3)}{k}. \tag{11}$$

This interval containing $\rho_k(X)$ is obtained by either (i) assigning the rank 1 to X for all $k-n$ members of the population not in the sample (lower bound for $\rho_k(X)$); or (ii) assigning the rank 3 to X for all $k-n$ non-sample population members (upper bound for $\rho_k(X)$). Taking $k = N$, a sample of size n provides a 100% confidence interval

$$\frac{N + n(\rho_n(X) - 1)}{N} \leq \rho_N(X) \leq \frac{3N + n(\rho_n(X) - 3)}{N}$$

for the population ranking $\rho_N(X)$ of any profile X.

Insight into the confidence interval (11) is gained when we write it in the form

$$\rho_n(X) - e_1 \leq \rho_k(X) \leq \rho_n(X) + e_3, \tag{12}$$

where e_1 is the maximum distance from $\rho_k(X)$ to $\rho_n(X)$ towards the lower ranking bound 1, and e_3 is the maximum distance from $\rho_k(X)$ to $\rho_n(X)$ towards the upper ranking bound 3. Note further that

$$\frac{k + n(\rho_n(X) - 1)}{k} = \rho_n(X) - e_1, \tag{13}$$

which implies

$$e_1 = \frac{k-n}{k}(\rho_n(X) - 1). \tag{14}$$

Similarly,

$$\frac{3k + n(\rho_n(X) - 3)}{k} = \rho_n(X) + e_2, \tag{15}$$

which yields

$$e_2 = \frac{k-n}{k}(3 - \rho_n(X)). \tag{16}$$

Let $\lambda = \frac{k-n}{k}$ be the proportion of the population which has not taken the survey. In both directions, the interval extends from $\rho_n(X)$ a distance λ times the distance to the endpoints of the ranking interval $[1, 3]$. In addition, the length of this confidence interval is given by $e_1 + e_2 = 2\lambda$, as seen in Figure 4. Note that the coefficient 2 of λ arises algebraically as the difference between the extreme rankings 1 and 3.

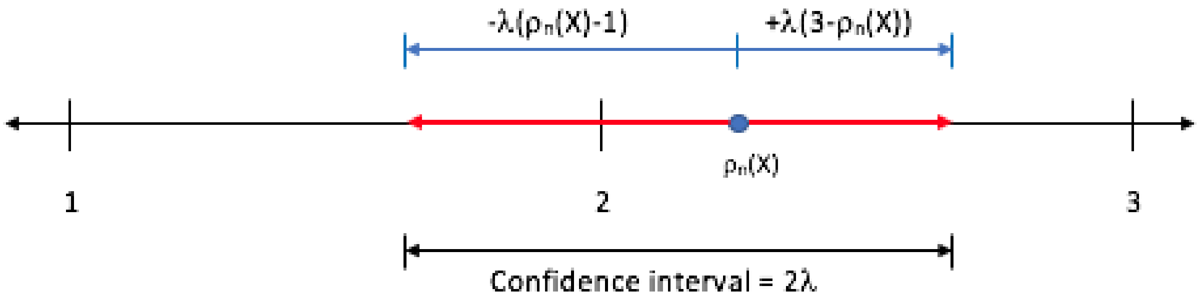

Figure 3: Given a profile X and its sample ranking $\rho_n(X)$, a 100% confidence interval for the population profile ranking $p_k(X)$ is determined by $\lambda = \frac{k-n}{k}$, the proportion of the population who have not taken the survey.

2.8 Application

These results can be used to quantify the possible consequences of survey response bias. For the simple example introduced in Section 2, let us consider the following response scenarios. Assume that, out of the total population $N = 8$, two respondents have outright refused to take the survey, four have completed it, and the other two have not yet replied. If the analysis is performed with only the sample of $n = 4$ respondents, then the length of the confidence interval is $\frac{2(8-4)}{8} = 1$. If one of the non-respondents is convinced to participate, the interval length is reduced to $\frac{2(8-5)}{8} = \frac{3}{4}$, improving the precision by 25%. If both of the non-respondents participate, then the interval is further reduced to $\frac{2(8-6)}{8} = \frac{1}{2}$. In other words, the two non-respondents cause the confidence interval to be twice as large, an important consideration in seeking to elicit survey response.

2.9 Attribute Decomposition Via 2 Dimensional Multi-mensional Scaling

2.10 Application: Profile Ranking Analysis of a Small Population Disaster Relief Survey

In this section we show how to apply PR ACBC methodology to an actual survey.

2.11 Humanitarian Disaster Relief

PR-based ACBC methodology for analyzing small populations has many possible applications. In disaster relief, effectiveness of a response may depend on the quality of collaboration between organizations with a broad diversity of religious and ideological perspectives. For effective coordination of relief, it is important that humanitarian organizations understand the unique traits and characteristics that shape their disaster response decisions. Through comparison of these factors, it is possible to design optimal partnerships and joint endeavors between organizations that may fulfill distinct, yet complimentary, humanitarian roles. Our research focused on how faith based organizations (FBOs) prioritize key attributes affecting their decision whether or not to respond to an international humanitarian disaster.

To this end, we designed a survey that creates "Go/No-Go" decision profile preferences for a small population of approximately 50 international FBOs with headquarters in the United States. The attributes and levels for this survey are displayed in Figure 4. Different disaster scenarios face off against each other in the tournament stage and are given rankings based on their performance. The purpose of the ranking is to determine whether FBOs fill a certain niche in disaster response landscape as might be inferred for by their "number-one" or "top-three" ranked disaster response profiles. As a case study, we discus how our methodology was applied to data obtained for this survey administered using Sawtooth's Lighthouse platform.

2.12 Ranking Method

As shown in Figure 4, our humanitarian survey consists of four attributes, each with three levels. Thus, the number of possible profiles is $3^4 = 81$. These are identified by four digit numbers $X = x_1x_2x_3x_4$ where profile X has level x_1 for the first attribute, level x_2 for the second attribute, level x_3 for the third attribute, and level x_4 for the fourth attribute. In the tournament stage of the competition, there are four rounds, in which sixteen profiles face off against each other in head to head match-ups, much like the FIFA World Cup Round of 16. The competing profiles are selected from the 81 possible profiles based on the respondent's BYO preferences. We assign a ranking of 1 to the tournament winner, 2 to the runner up, 3 to the semifinal losers, 4 to the quarterfinal losers, 5 to the profiles that are eliminated in the first round, and 6 to those that do not appear in the tournament. By the two-week deadline after deploying the survey, 5 FBOs had responded, resulting in the sample ranking table shown in Table 11. The responding population ($N = 10$) most likely to produce this observed sample is also included.

Profile	1	2	3	4	5	6	7	8	9	10	11	12	13	PR
1212	2	3	7	3	7	7	1	7	7	7	4	1	7	4.846
1112	7	1	7	7	1	2	7	5	7	3	7	7	4	5.000
1312	7	7	7	7	5	5	7	1	1	1	7	7	3	5.000
3212	7	7	7	5	5	1	2	7	4	7	7	7	1	5.154
1122	1	7	7	4	7	7	5	7	4	7	3	3	7	5.308
2212	7	5	5	7	7	7	7	2	7	4	2	2	7	5.308
1222	7	7	4	7	4	5	7	4	3	2	7	7	7	5.462
1232	5	7	7	5	5	5	4	5	7	5	7	7	2	5.462
2312	7	7	7	1	5	3	4	7	2	7	7	7	7	5.462
2112	3	7	7	4	5	4	3	7	7	7	7	7	5	5.615

Table 11: Top 10 ranked profiles for FBOs ($n = 13, N = 50$)

2.13 Profile Rank Confidence Intervals

Applying the results of Section 3, using the data in Table 11, given a profile's sample ranking ($n = 5$), we can construct confidence intervals for population

What is the focus of this survey?

This survey asks you to identify the type of international disaster relief scenarios your organization would respond to. Your decision may be based either on prior experience, or on the degree to which the scenario alligns with your organization's core values. With input from several managers of international relief organizations, we have framed your "go/no go" decision according to the following factors:

Factor	Description	Levels Considered
EXTERNAL FUNDING (Unrestricted money donated for the response.)	What amount of external funding is available for your response?	a) At least 75% of what is required b) About 50% of what is required c) Less than 25% of what is required
RESPONSE SCALE	Has the disaster been declared an IASC Level 1,2 or 3 emergency?	a) Large INGOs are responding in clusters (Level 3) b) Some INGOs are responding (Level 2) c) Few or no INGOs are responding (Level 1 or undeclared)
NEED ASSESSMENT	What is the level of need of the vulnerable for your organization's resources and/or capabilities?	a) Clear need for your organization's particular contribution b) Contribution could be of assistance but is not indispensable c) Unknown need of contribution
ACCESS TO AFFECTED COMMUNITY	What partnership provides connection to the community?	a) No pre-existing partnership is available b) A local partner has requested help c) An outside party has invited participation

Figure 4: An ACBC survey with 4 attributes consisting of 3 levels each.

rankings with $N = 10$. For example, consider the profile 1232 which had the top ranking in the sample. In a population with $N = 10$ and $\rho_n(1232) = 4.2$, the value $\lambda = \frac{1}{2}$ results in the confidence interval shown in Figure 5. Since the range of individual rankings is $6 - 1 = 5$, we compute the interval length as $5\lambda = 2.5$. It follows that that the population ranking $\rho_k(2113)$ could be up to 1.6 less or .9 greater than the sample ranking $\rho_5(1232) = 4.2$. After following up with organizations to whom we sent the survey, we received survey data from an additional 5 FBOs, and calculated $\rho_{10}(1232) = 4.6$, which falls within our confidence interval (Figure 5). This approach can be applied to any profile in Table 11 and provides a simple visualization of the extent to which sample profile rankings can capture population profile rankings.

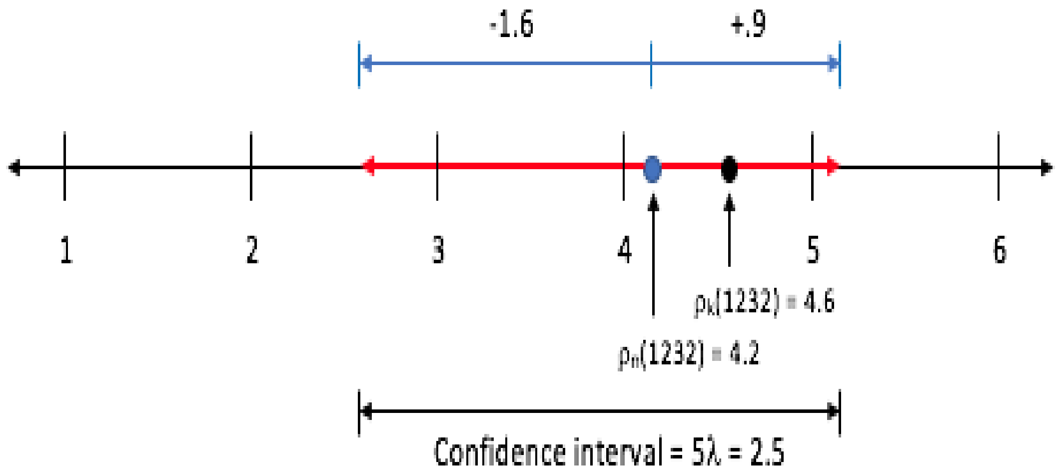

Figure 5: Confidence Interval for profile 1232 (n = 5, k = 10).

2.14 Visualization of Top Ranked Profiles

A final outcome of PR ACBC would be a visual representation of the top ranked profiles. For the population rankings in Table **??**, there is a triple tie for the number 1 ranking. A visual display of top-ranked profiles such as shown in Figure 6, might enhance the GUI of existing ACBC software.

3 Conclusion

Unlike conjoint analysis of survey data where the target populations are large and more suitable for conventional statistical tools, we have introduced a simple, intuitive approach to a small population's profile rankings based on sample data. The population most likely to yield the given sample results are expected to have the same profile rankings as the sample's. We also provide a new type of 100% confidence interval for profile rankings without using standard deviations, which can be used to quantify maximum possible survey response bias. Furthermore,

Figure 6: A visual lineup of top-ranked profiles could enhance the GUI of commercial ACBC software.

we have shown that part-worth utilities obtained by multiple linear regression can only replicate sample profile rankings under special conditions, with the residuals indicating the errors in predicted rankings.

For applications such as humanitarian disaster relief, sample profile ranks are more easily understandable to respondents than part-worth utilities. The intuitive nature of profile ranking allows quick and straightforward analysis of any small sample of disaster response organizations. Given that the sample comprises a relatively significant portion of the overall population, 100% confidence intervals provide an absolute range of possible population profile rankings without the complexity of statistical inference. Consequently, for small organizations with low operating costs, these techniques may serve as an effective, yet affordable, alternative to more sophisticated ACBC survey utility-based analysis software. Moreover, a visualization of top-ranked or bottom-ranked profiles is a good way to summarize PR ACBC survey results.

Major areas open to further research include analysis of different ranking systems for various types of choice tournaments and application of PR ACBC methodology to other small population studies.

Acknowledgements

The authors would like to thank Erica Gralla, Jarrod Goenzel, Timotius Kartawijaya, and Mike Veatch for their valuable contributions to this work.

References

Alvo, M., and Yu, P.L.H. 2014. *Statistical Methods for Ranking Data*. Springer.

Gralla, E., Goentzel, J., and Fine G. 2014. Assessing trade-offs among multiple objectives for humanitarian aid delivery using expert preferences. *Production and Operations Management*, Springer-Verlag Berlin 23(6), 978-989.

Oberhofer, W. and Kaufman, H. Maximum Likelihood Estimation of a Multivariate Hypergeometric Distribution. *Sankhya: The Indian Journal of Statistics, Series B (1960-2002)*, Indian Statistical Institute, 49(2), 188-191.

Orme, B.K., and Chrzan, K. 2017. *Becoming an Expert in Conjoint Analysis: Choice Modeling for Pros.* Sawtooth Software.

Rao, V. R. 2014. *Applied Conjoint Analysis.* Springer.

Stewart, J. 2016. *Calculus, Early Transcendentals (8E).* Cengage Learning.

Rossi, P., Allenby, G. and McCulloch R. 2005. *Baysian Statistics and Marketing.* John Wiley & Sons, Ltd.