

## データ解析環境Rによる選択型コンジョイント分析入門

合崎英男<sup>\*</sup>・西村和志<sup>\*\*</sup>

目 次	
緒 言 .....	151
例題の設定 .....	152
1 評価シナリオ .....	152
2 回答結果データ .....	153
3 確定効用関数の設定 .....	153
AlgDesignパッケージを用いた全選択肢集合の作成 .....	153
1 全選択肢集合の作成手順 .....	153
2 ステップ1: 完全実施要因計画の作成 .....	154
3 ステップ2: 実験回数の削減による初期集合の作成 .....	155
4 ステップ3: 初期集合からのコピー集合の作成 .....	156
5 ステップ4~6: 非復元無作為抽出による選択肢集合の作成と質問の作成 .....	157
survivalパッケージを用いたデータ解析 .....	158
1 説明変数が属性変数のみのケース .....	159
2 説明変数に個人・世帯特性が含まれるケース .....	162
結 言 .....	163
補 論 .....	164
1 実験回数について .....	164
2 計測結果の活用について .....	165
参考文献 .....	166
Summary .....	168
付 録 .....	169

## 緒 言

近年、新規に開発された農業技術の経営評価や複数の環境要因を考慮した環境経済評価など、幅広い分野で選択型コンジョイント分析が活用されてきている（合崎2005）。しかし、選択型コンジョイント分析を実施するためには、実験計画法を活用して選択肢集合を作成する作業（質問設計）と得られたデータを離散選択モデルにより統計分析する作業が必要とされるため、はじめて活用しようとする人にはいくぶん敷居の高い手法となっている。加えて、それらの作業を一貫した日本語環境で容易に行えるソフトウェアが見当たらないことも、活用しづらい状況を作り出している。

そこで本稿では、データ解析環境R（R Development Core Team（2006）：以下、R）を用いた選択肢集合の作成とデータ分析の手順について解説する。Rはさまざまな統計手法を実行することができ、標準的な表計算ソフトで描ける以上の作図機能が実装されている。また、自由に使うことができ、インターネットへの接続環境が整

っている人にとっては、ほとんど追加的な費用をかけずに入手できる。さらに、GUI環境や主要なメッセージを日本語化することができ、計量経済学向けのソフトウェアにみられる英語表示によるとつきにくさも軽減されている。

ただし、プログラムを実行するためには基本的にコマンドを入力、実行する必要があるが、操作性という点では既存の計量経済学向けのソフトウェアと同等、あるいはそれ以下となっている。しかし、本稿が紹介する手順で選択型コンジョイント分析を行うためには、主にAlgDesignとsurvivalという2つのパッケージに含まれる少数のコマンドを使用するだけである。全体を通じて手順を体験すれば、必要な知識は習得できよう。本稿末には付録として使用するデータとコマンドの一覧を掲載している。実際にRを動かしながら手順を習得されたい。なお、計測結果の具体的な活用例（属性の価格評価と選択確率の予測）については、補論にまとめているので参考にされたい。

本稿の読者としては、選択型コンジョイント分析を実施するための基本的な手順は一通り学習しているが、具体的にはどのようなソフトウェアをどのように使えばいいのかかわからない人を前提としている。本稿では紙幅の関係から選択型コンジョイント分析の概要等は省略している。手法についての知識を得たい方は既存文献を参照されたい（たとえば、合崎2005、栗山・庄子2005、Louviere et al. 2000、Benett and Blaney 2001、Champ et

<sup>\*</sup> 農村計画部事業評価研究室

<sup>\*\*</sup> 九州沖縄農業研究センターイネ発酵TMR研究チーム

平成19年3月12日受理

キーワード：R、コンジョイント分析、選択実験、実験計画、選択肢集合、離散選択モデル、条件付きロジットモデル

al. 2003, Hensher et al. 2005)。また、本稿ではR本体や後述する必要なパッケージがインストールされていることを前提としている。このようなRやパッケージのインストール方法、基本的な使い方について知りたい方は、Rの既存文献を参照されたい(たとえば、岡田2004、船尾2005、船尾・高波2005)。日本語によるR関連の最新の情報についてはRjpWiki (<http://www.okada.jp.org/RWiki/index.php?RjpWiki>) から入手できるので、関心のある方は参照されたい。

なお、本稿に掲載されている情報を利用して何らかの損失が生じたとしても、その原因がいかなるものであっても、Rの著者や筆者らは責任を負わない。自己責任のもとで利用されたい\*1。

本稿の作成にあたっては、多くの方々のご協力を得た。人数の多さと紙幅の関係からお一人ずつお名前を挙げることはできないが、この場を借りて感謝申し上げる。

### 例題の設定

#### 1 評価シナリオ

読者ができるだけ容易に手順を理解できるように、選択型コンジョイント分析の適用事例を1つ想定して解説を行う。具体的には、Fig.1に示す形式の質問を通じて、農家による堆肥の購入行動を選択型コンジョイント分析

\*1 筆者らはRのエキスパートではないので、本稿で紹介している手順に比べて、より効率的に目的を達せられる方法があるかもしれない。お気づきの点があればご連絡いただきたい。

で分析する状況を考える。どのような堆肥をいくらで販売すれば、農家は堆肥を購入するか(どの程度の確率で購入するか)を明らかにする課題に選択型コンジョイント分析を適用しようとしている状況である。

堆肥を構成する属性は、「1トン当たり堆肥価格」、「堆肥の運搬・散布サービス」、「土壌診断・施肥設計サービス」である。「堆肥の運搬・散布サービス」は、購入した堆肥を指定した圃場(農地)まで運搬し、散布してくれるサービスである。「土壌診断・施肥設計サービス」とは、堆肥を投入する圃場の肥料成分等の土壌診断を実施し、その結果を踏まえてどの程度堆肥等を投入すればよいか助言してくれるサービスである。これら2種類のサービスは、それぞれ「あり」と「なし」の2水準である。価格については、サービスが付加される堆肥であれば、その対価も含まれる。価格の水準は、「1,000円」「2,000円」「3,000円」「4,000円」の4つとする。以上の3つの属性以外の点(たとえば、堆肥の肥料成分等の品質)については、堆肥Aと堆肥Bでは同等と仮定する。

質問では、Fig.1に示すように各水準の組み合わせから構成される2つの堆肥Aと堆肥Bが提示される。回答農家は、どちらかを購入するか、あるいは「どちらも買わない」の3つの選択肢のうちの1つを選ぶ意思決定を行うことになる。

本稿ではFig.1形式の質問が全体で8問あり(全選択肢集合が8問から構成されている)、各農家はそれら8問すべてに回答すると設定する。具体的な組み合わせは、Table 1に示すとおりである。1行が1つの質問で提示さ

問1 つぎの2つの堆肥が販売されています。どちらを購入したいですか(1つに○)。

1. 堆肥A	2. 堆肥B	3. どちらも買わない								
1 tあたり堆肥価格 堆肥の運搬・散布 土壌診断・施肥設計	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="text-align: center;">堆肥A</td></tr> <tr><td style="text-align: center;">4000円</td></tr> <tr><td style="text-align: center;">あり</td></tr> <tr><td style="text-align: center;">あり</td></tr> </table>	堆肥A	4000円	あり	あり	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="text-align: center;">堆肥B</td></tr> <tr><td style="text-align: center;">2000円</td></tr> <tr><td style="text-align: center;">なし</td></tr> <tr><td style="text-align: center;">あり</td></tr> </table>	堆肥B	2000円	なし	あり
堆肥A										
4000円										
あり										
あり										
堆肥B										
2000円										
なし										
あり										
	どちらも買わない									

Fig.1 本稿で想定する選択型コンジョイント分析の質問例

An example of a questionnaire for choice experiments

Table 1 本稿で使用する全選択肢集合(「どちらも買わない」を除く)

Choice sets, excluding the none of these option

質問	堆肥A			堆肥B		
	PRICE	TRANS	SOIL	PRICE	TRANS	SOIL
1	4000	1	1	2000	0	1
2	1000	0	0	1000	1	1
3	2000	0	1	4000	0	0
4	3000	0	1	1000	0	0
5	3000	1	0	3000	0	1
6	1000	1	1	4000	1	1
7	2000	1	0	3000	1	0
8	4000	0	0	2000	1	0

れる 2 つの堆肥の内容を示している。たとえば、1 行目は問 1 で提示される堆肥Aと堆肥Bを示しており、堆肥Aは価格 (PRICE) が4,000円、運搬・散布サービス (TRANS) が「あり (1)」, 土壌診断・施肥設計サービス (SOIL) が「あり (1)」である。同様に、堆肥Bは価格が2,000円、運搬・散布サービスが「なし (0)」, 土壌診断・施肥設計サービスが「あり (1)」であることがわかる。

後述するように、本稿ではRのAlgDesignパッケージを活用しながら、Table 1に示す 8 問分の全選択肢集合を作成する。

## 2 回答結果データ

Fig.1に例示した形式による全 8 問 (Table 1) に10名の農家が回答したとして、サンプルサイズ (観測値数) を80 (= 10名 × 8問) と設定した<sup>\*2</sup>。回答結果データは、人工的に作成した (栗山2000)。また、回答者 (農家) の個人・世帯特性 (経営特性) が各属性の評価、したがって回答パターンに与える影響を検証する手順を示すため、経営耕地面積の違いが堆肥の運搬・散布サービスと土壌診断・施肥設計サービスの評価に与える影響を検討することとした。使用するデータセットについては、本稿末に付録として掲載しているのを参照されたい<sup>\*3</sup>。

## 3 確定効用関数の設定

本稿では、選択型コンジョイント分析の適用事例で基本モデルとして最も活用されている条件付きロジット・モデル (conditional logit model) によりデータの分析を行う。

Fig.1に示した選択肢集合において、「どちらも買わない」選択肢の確定効用 $V_C$ を基準 ( $V_C=0$ ) として、堆肥Aと堆肥Bの各確定効用 $V_A$ と $V_B$ を、以下のように設定する。

$$V_A = ASC + B_P PRICE_A + B_T TRANS_A + B_S SOIL_A$$

$$V_B = ASC + B_P PRICE_B + B_T TRANS_B + B_S SOIL_B$$

ただし、ASCは選択肢固有定数、 $PRICE_i$  ( $i=A,B$ ) と $B_P$ は堆肥  $i$  の価格 (単位: 円/トン) とその係数、

\*2 本稿では例示のため可能な限りサンプルサイズを小さくすることを意図している。実際に選択型コンジョイント分析を適用するには、10名の回答者のデータから有益な情報が引き出せることはほとんどあり得ないだろう。ここでのサンプルサイズが、実証分析を行う際のサンプルサイズの下限であると誤解してはならない。

\*3 付録には、本稿で入力するコマンドも一括して掲載している。なお、これまで通りであれば、本稿のPDF版が農村工学研究所のウェブサイト (<http://rkk.naro.affrc.go.jp>) 内にアップロードされる。PDF版を入手すれば、データ部分をテキストとして抽出することが可能なはずなので、手作業で入力する手間を省くことができるであろう。

$TRANS_i$ と $B_T$ は堆肥の運搬・散布サービスの有無を表すダミー変数 (1:あり, 0:なし) とその係数、 $SOIL_i$ と $B_S$ は土壌診断・施肥設計サービスの有無を表すダミー変数 (1:あり, 0:なし) とその係数、である。堆肥Aと堆肥Bは、価格と2つのサービスの有無以外の点では同等と仮定する。そのため、両堆肥で価格やサービスに対する農家の評価が変化するとは考えられず、ここでは2つの堆肥で共通する変数の係数は等しいという制約を課す。たとえば、堆肥の価格は堆肥Aでは変数 $PRICE_A$ 、堆肥Bでは $PRICE_B$ と表されているが、それらの係数はどちらも同じ $B_P$ となる。

## AlgDesignパッケージを用いた全選択肢集合の作成

### 1 全選択肢集合の作成手順

本稿では、Louviere et al. (2000, p.114) が紹介している全選択肢集合の作成方法のうち、次の手順により全選択肢集合を作成する (Fig.2)<sup>\*4</sup>。全選択肢集合の作成では、RのAlgDesignパッケージ (Wheeler 2006) を用いる。なお、以下では「どちらも買わない」等の「非選択」選択肢を除いて1つの質問で提示される選択肢数を $M$ 、選択型コンジョイント分析用の全質問数を $P$ としている。

ステップ1: 完全実施要因計画を作成する。

ステップ2: 完全実施要因計画の実験回数を削減 (実験を一部削除) して初期集合を作成する。

ステップ3: 初期集合のコピーを $M-1$ 組作成する。

ステップ4: 初期集合とコピーを合わせた $M$ 組の選択肢集合から、それぞれ非復元無作為抽出により1つの選択肢を取り出す。これにより、選択型コンジョイント分析の質問1つで提示する選択肢集合ができあがる。したがって、1つの質問では («どちらも買わない」といった非選択選択肢を除いて)  $M$ 個の選択肢が提示される。

ステップ5: 上記ステップ4の作業を計 $P$ 回繰り返す。その際、同じ水準から構成される選択肢が1つの質問で並ぶことのないようにする。同じ選択肢が現れたら抽出作業をやり直す。

ステップ6: 得られた $P$ 問分の選択肢集合に対して、それぞれ「どちらも買わない」選択肢を追加し、さらに各属性に具体的な水準を割り付けることで、最終的に調査票で提示する選択型コンジョイント分析用の質問を作成する。

\*4 日本語による紹介には柘植ら (2005) がある。

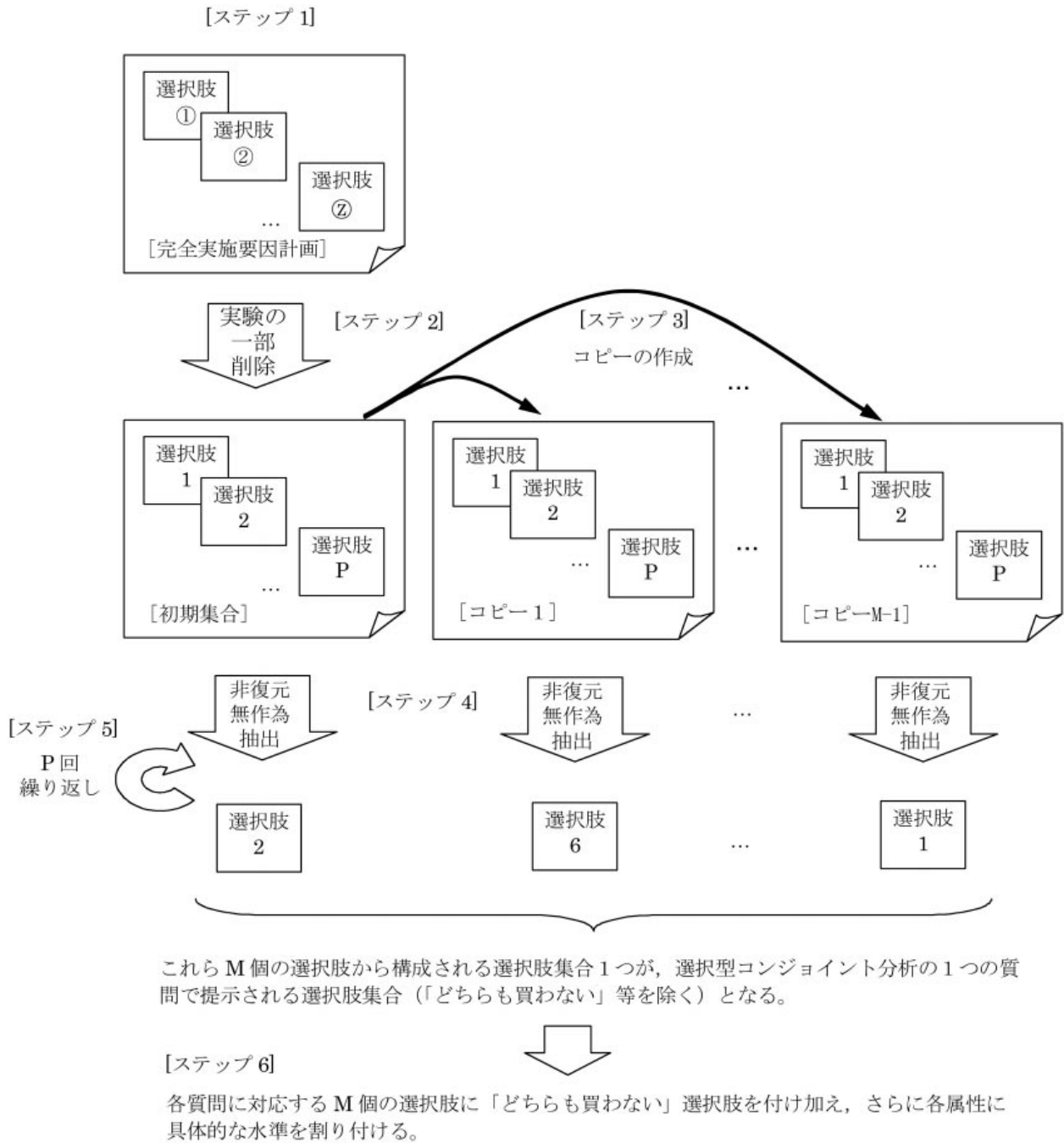


Fig.2 全選択肢集合（「どちらも買わない」を除く）の作成手順  
 Steps in constructing choice sets, excluding the none of these option

2 ステップ 1：完全実施要因計画の作成

AlgDesignパッケージに含まれる関数gen.factorial()を使って、価格に対応する 4 水準系の要因が 1 つと、2 種類のサービスに対応する 2 水準系の要因が 2 つから構成される完全実施要因計画を作成する。完全実施要因計画とは、使用するすべての要因の全組み合わせである。ここでは 4 水準系の要因が 1 つと 2 水準系の要因が 2 つであることから、 $4 \times 2 \times 2 = 16$ 通りが完全実施要因計画としての全組み合わせ数となる。関数library()を使ってAlgDesignパッケージを呼び出した上で、本例では次のようにコマンドを入力、実行する（Enter キーを押す）

ことで、目的とする完全実施要因計画が得られる（Fig.3の 1 行目と同じ）。なお、本稿では入力するコマンドは改行することなく 1 行で入力するが、本文では書式的関係から 2 行以上にわたって記述することがある。正しいコマンドの入力に対応する図を参照されたい。

```
ffd<-gen.factorial(c(4,2,2),varNames=c("PRICE",
"TRANS","SOIL"),factors="all")
```

ここで、右辺（「<」の右側）の関数gen.factorial()が、完全実施要因計画を作成するコマンドである。作成され



```

> ffd<-gen.factorial(c(4, 2, 2), varNames=c("PRICE", "TRANS", "SOIL"), factors="all")
> ffd
  PRICE TRANS SOIL
1     1     1     1
2     2     1     1
3     3     1     1
4     4     1     1
5     1     2     1
6     2     2     1
7     3     2     1
8     4     2     1
9     1     1     2
10    2     1     2
11    3     1     2
12    4     1     2
13    1     2     2
14    2     2     2
15    3     2     2
16    4     2     2

```

Fig.3 関数gen.factorial()による $4^1 2^2$ 完全実施要因計画

A full factorial design with one 4-level factor and two 2-level factors generated from the function gen.factorial()

た完全実施要因計画は、列が要因に、行が実験に対応する行列形式となる（Fig.3の3行目以降：詳細は後述）。

括弧内にある「c(4,2,2)」は4, 2, 2という数値の入ったベクトルを作成するコマンドである。作成しなければならない計画が4水準系の要因（属性）が1つ、2水準系の要因が2つであることから、左側の列から順に4水準系の要因、2水準系の要因、2水準系の要因という3列として作成するよう指示している。したがって、作成しようとする完全実施要因計画によって、ここで指定する値は変化する。たとえば、6水準系の要因が1つ、4水準系の要因が2つ、2水準系の要因が3つとすれば、「c(6,4,4,2,2,2)」となる。

つぎの「varNames=c("PRICE", "TRANS", "SOIL")」は、作成した完全実施要因計画に含まれる各要因（列）に与える名称を指定している。左側の列から順に、PRICE（堆肥価格）、TRANS（堆肥の運搬・散布サービス）、SOIL（土壌診断・施肥設計サービス）と命名している。名称を指定しないと、X1, X2, X3等の一般的な名称が自動的に割り当てられる。

最後の「factors="all"」は、指定したすべての要因（今回は4水準系が1要因と2水準系が2要因の計3要因）を使って完全実施要因計画を作成するよう指示しており、本稿の手順に沿っている限りは、このように入力すればよい。

以上のように構成されるコマンドを実行すると、 $4^1 2^2$ の完全実施要因計画がデータフレームffdに作成される（「<-」は右辺（の結果）を左辺に代入することを指示）。ここでRにデータフレーム名

```
ffd
```

と入力、実行すれば、全体で実験回数16回の完全実施要因計画がデータフレームffdに作成されたことを確認できる（Fig.3の2行目以降）。ここでいう「実験」1回が、選択型コンジョイント分析における「選択肢（プロファイル）」1つに対応する。

なお、作成された完全実施要因計画を格納するデータフレーム名をffdとしたが、この名前は任意である。以下でもデータフレームや変数の名称を新たに指定するが、いずれの場合もその名称は任意である。ちなみに、日本語化したRでは、データフレーム名や変数名に漢字等を使うことができる。

### 3 ステップ2：実験回数の削減による初期集合の作成

完全実施要因計画では、一般に実験回数が多くなる。そこで、AlgDesignパッケージに含まれる関数optFederov()を使って実験回数を削減し、初期集合を作成する。今回は、実験回数16回を半分の8回にする1/2部分実施要因計画とする。そのためには、次のようにコマンドを入力、実行する。

```
des<-optFederov(~.,ffd,8)
```

関数optFederov()の括弧内のうち、「ffd」は実験回数の削減対象となる元の完全実施要因計画が保存されているデータフレーム名、「8」は目標とする部分実施要因計画の実験回数をそれぞれ表す。「~.」は、データフレームffdに含まれる要因名を用いて、すべての要因が線型モデルとして処理されることを指定している。別のモデルを仮定することもできるが、基本的には「~.」に設定するものと理解すればよい<sup>\*5</sup>。

```

> runif(1)
[1] 0.01887456
> des<-optFederov(~., ffd, 8)
> des
$D
[1] 0.25

$A
[1] 6.333333

$Ge
[1] 1

$Dea
[1] 1

$design
  PRICE TRANS SOIL
1      1      1  1
4      4      1  1
6      2      2  1
7      3      2  1
10     2      1  2
11     3      1  2
13     1      2  2
16     4      2  2

$rows
[1] 1 4 6 7 10 11 13 16

> cs1<-des$design
> cs1
  PRICE TRANS SOIL
1      1      1  1
4      4      1  1
6      2      2  1
7      3      2  1
10     2      1  2
11     3      1  2
13     1      2  2
16     4      2  2

```

Fig.4 関数optFederov()による1/2部分実施要因計画の作成

A half-fractional factorial design generated from the function  
optFederov()

ここで1つ注意が必要である。関数optFederov()を実行する前に乱数を一度発生させる必要がある。たとえば、

```
runif(1)
```

と入力、実行し、一様乱数を1つ生成する。この作業を

\*5 記号「~」はtilde (チルダ, ティルダ)であり, 最も一般的な106 (109)型キーボードであれば, 半角入力状態で「シフト」キーを押しながら「^ (ハット)」キー (キーボード右上の「¥」キーの左, あるいは数字の「0 (わ)」キーの2つ右にあるキー)を押すことで入力する。

```

> cs2<-cs1
> cs2
  PRICE TRANS SOIL
1      1      1  1
4      4      1  1
6      2      2  1
7      3      2  1
10     2      1  2
11     3      1  2
13     1      2  2
16     4      2  2

```

Fig.5 初期集合のコピー作成

Making a copy of the initial choice sets

行わないと, 関数optFederov()を上記のように実行しても, Random.seedが存在しないとしてエラーが生じる\*6。

ここではFig.4のように関数runif()を実行した上で, 関数optFederov()を使ったコマンドを実行する。その結果が保存されているオブジェクト名

```
des
```

を入力, 実行することで結果を確認できる\*7。その中には, 求めたい部分実施要因計画以外の情報も含まれている。そこで, 作業目的である1/2部分実施要因計画だけをデータフレームcs1に取り出すため,

```
cs1<-des$design
```

と入力, 実行する。

なお, ここでは部分実施要因計画の実験回数を8回としたが, これをどの程度とするかについては補論1を参照されたい。

#### 4 ステップ3: 初期集合からのコピー集合の作成

ステップ3では, 初期集合 (今回であればデータフレームcs1)を必要な分だけコピーする。今回は, 提示される堆肥が2種類であることからM=2, したがってコピーを1つ作成する。コピーとなる集合のデータフレーム名をcs2とすれば, 次のコマンドを入力, 実行することにより, データフレームcs1から簡単にコピーを作成できる (Fig.5)。

```
cs2<-cs1
```

\*6 乱数を発生させる代わりに, 関数set.seed()でrandom seedを指定する方法もある。

\*7 関数optFederov()は部分実施要因計画の作成に際して乱数を利用しているため, 同じコマンドを入力しても同じ計画や相関係数が得られるとは限らない。関数optFederov()が使用するアルゴリズムについては, AlgDesignパッケージに含まれるドキュメント (Wheeler 2004)を参照されたい。

ここでコピーにより作成されたデータフレーム名

cs2

を入力，実行すれば，データフレーム<sub>cs1</sub>に格納されている部分実施要因計画（Fig.4）と同じ計画がデータフレーム<sub>cs2</sub>に保存されていることを確認できる。

もしも，選択型コンジョイント分析の質問（選択肢集合）1つで提示される堆肥が3つであるとし，3つめの選択肢を作成するための集合をデータフレーム<sub>cs3</sub>と命名すれば，

cs3<-cs1

とすればよい。1つの質問で提示する堆肥がM個の場合，同様な手順で計M-1個のコピーを作成する。

5 ステップ4～6：非復元無作為抽出による選択肢集合の作成と質問の作成

ステップ4は，初期集合を含めた各選択肢の集合から非復元抽出により無作為に選択肢を1つ取り出す作業である。Rで実施するときには，非復元抽出作業を計P回繰り返すというステップ5の作業と同時に進行。具体的には，上記で作成した各選択肢の集合に対して同じ行数

分の乱数を発生させ，その乱数を使って各選択肢の集合内で選択肢を並び替える。

まず，データフレーム<sub>cs1</sub>に変数<sub>r1</sub>と命名した乱数を追加する（Fig.6）。発生させる乱数の個数はデータフレーム<sub>cs1</sub>の行数と一致させる必要があるので，ここでは8つとなる。具体的には，下記のコマンドを入力，実行することで必要な乱数を変数<sub>r1</sub>としてデータフレーム<sub>cs1</sub>に加える。

cs1<-transform(cs1,r1=runif(8))

ここで，右辺の「r1=runif(8)」は，変数<sub>r1</sub>に一樣乱数を8個生成，保存するという命令である。「transform(cs1, r1=runif(8))」は，変数<sub>r1</sub>をデータフレーム<sub>cs1</sub>に追加するという命令である。たとえば，データフレーム<sub>cs1</sub>がn行から構成されているとすれば，上記のコマンドのうち，「runif(8)」の「8」の部分「n」に変更して「runif(n)」とする。

同様に，データフレーム<sub>cs2</sub>についても，次のコマンドを入力，実行して，変数<sub>r2</sub>で表される一樣乱数を追加する。

cs2<-transform(cs2,r2=runif(8))

全体でM個のデータフレーム（選択肢の集合）がある場合は，同様な作業を計M回繰り返すことになる。

つぎの作業は，各データフレームに含まれる選択肢を，上記で生成した乱数にもとづいて並び替える作業である。データフレーム<sub>cs1</sub>については，baseパッケージ（Rと同時に読み込み済み）に含まれる関数`order()`を使って，次のコマンドを入力，実行する。

cs1\_sort<-cs1[order(cs1\$r1),]

ここで，右辺の「order(cs1\$r1)」は，データフレーム<sub>cs1</sub>に含まれる変数<sub>r1</sub>に基づいてデータを並び替えた

```
> cs1<-transform(cs1, r1=runif(8))
> cs1
  PRICE TRANS SOIL      r1
1     1     1     1 0.07809708
4     4     1     1 0.94791114
6     2     2     1 0.86082508
7     3     2     1 0.43025533
10    2     1     2 0.25497669
11    3     1     2 0.41863009
13    1     2     2 0.72482487
16    4     2     2 0.04784136
```

Fig.6 データフレーム<sub>cs1</sub>への変数<sub>r1</sub>の追加  
Adding a new variable r1 to a data frame cs1

```
> cs1
  PRICE TRANS SOIL      r1
1     1     1     1 0.07809708
4     4     1     1 0.94791114
6     2     2     1 0.86082508
7     3     2     1 0.43025533
10    2     1     2 0.25497669
11    3     1     2 0.41863009
13    1     2     2 0.72482487
16    4     2     2 0.04784136
```

```
> cs1_sort1<-cs1[order(cs1$r1),]
> cs1_sort
  PRICE TRANS SOIL      r1
16    4     2     2 0.04784136
1     1     1     1 0.07809708
10    2     1     2 0.25497669
11    3     1     2 0.41863009
7     3     2     1 0.43025533
13    1     2     2 0.72482487
6     2     2     1 0.86082508
4     4     1     1 0.94791114
```

Fig.7 並び替え前後のデータフレーム<sub>cs1</sub>の比較  
Comparing data frames cs1 and cs1\_sort

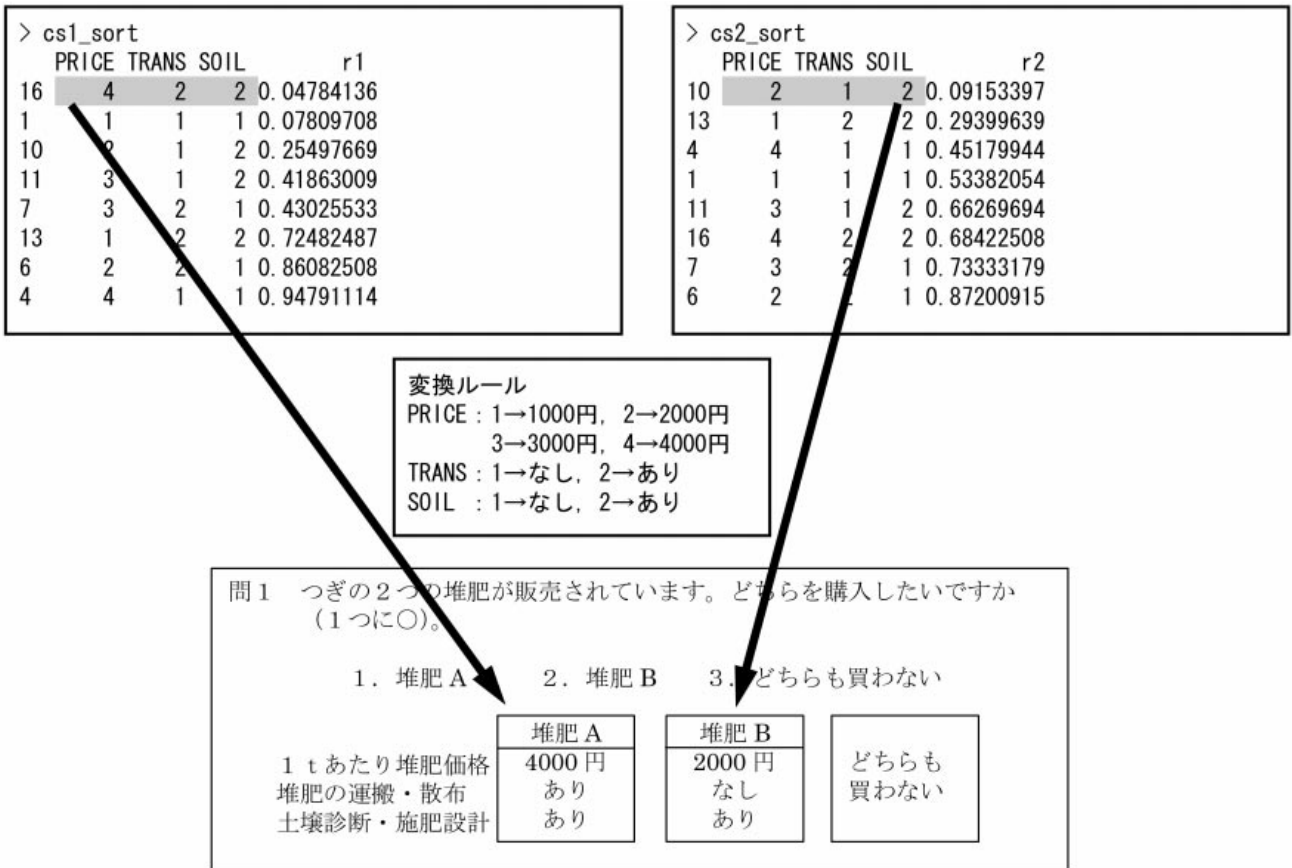


Fig.8 作成したソート後の選択肢集合と選択型コンジョイント分析の質問形式への変換  
Creating a choice experiments survey using data frames cs1\_sort and cs2\_sort

きの行番号の順番ベクトルを生成するという命令である。その順番ベクトルにしたがってデータフレームcs1に含まれる各変数を行単位で並び替えて(順番ベクトルに応じて各変数を呼び出す), cs1\_sortという新しいデータフレームに格納する,というのが上記のコマンドの意味である。この結果は,変数 r1を利用した並び替える前のデータフレームcs1と並び替えた後のデータフレームcs1\_sortを並べてみれば一目瞭然である (Fig.7)。

同様な作業を,初期集合からのコピーとして作成したすべてのデータフレームに対して実施する。ここでは,コピーはデータフレームcs2の1つだけなので,下記のコマンドを1回実施する。

```
cs2_sort<-cs2[order(cs2$r2),]
```

並び替えた結果(データフレームcs1\_sortやcs2\_sort)の各行が,上から順に選択型コンジョイント分析の各質問となる (Fig.8)。たとえば,データフレームcs1\_sortの1行目が問1で提示される堆肥Aの水準の組み合わせ,2行目が問2で提示される堆肥Aの水準の組み合わせ,8行目が問8で提示される堆肥Aの水準の組み合わせとなる。同様に,データフレームcs2\_sortの1行目が問1の堆肥Bの水準の組み合わせ,8行目は問8の堆肥Bの水準の組み合わせとなる。

そして,同じ順位の2つの選択肢が1つの選択型コンジョイント分析の質問で提示される2つの堆肥となり,それに「どちらも買わない」を追加することで,回答者に提示する選択型コンジョイント分析向けの質問が完成する(ステップ6)。

なお,同じ順位の2つの選択肢(同じ質問として提示される2つの選択肢)が同じ水準の組み合わせであるときには,一様乱数の発生作業(ステップ4)からやり直し(同じコマンドを入力,実行する),同じ選択肢が並びないようにする\*8。

### survivalパッケージを用いたデータ解析

選択型コンジョイント分析形式のデータを分析するためには離散選択モデルを用いる。離散選択モデルにはさまざまな種類のモデルが含まれるが,本事例では最も基本的な条件付きロジット・モデル (conditional logit model) を用いる。Rではsurvivalパッケージに含まれる

\*8 一様乱数の発生作業からやり直す場合,コマンドを再入力する必要はない。Rでは,キーボードの矢印キーの上( )を押すことで,以前に入力したコマンドが再表示される(詳しくは,たとえば船尾(2005)のp.23を参照)。この機能を使って必要なコマンドを再表示させればよい。



関数 `cllogit()` (Lumley 2006) が該当する。

選択型コンジョイント分析では、回答者の個人・世帯特性（以下、個人特性）が回答結果に与える影響を明示的に考慮するために、個人特性と属性変数との交差項（個人特性変数と属性変数の積によって表される変数）を含めてモデルを構築する場合がある。本稿でも、属性変数のみで定式化したモデルと、属性変数と個人・世帯特性により定式化したモデルの2つについて解説する。

以下、最初に説明変数が属性変数のみで定式化されたモデルについて、データセットの準備から計測結果の見方まで解説する。ついで、属性変数と個人特性変数の両方を説明変数として組み込んだモデルについて、データセットの作成方法と分析手順について解説する。

## 1 説明変数が属性変数のみのケース

### a データセットの作成と保存

関数 `cllogit()` で分析するデータセットは表計算ソフトなどで作成し、テキスト形式で保存した後、Rに読み込む。今回はCドライブの直下に `test` ディレクトリを作成し (`C:\test`)、ファイル名を「`data1.txt`」として `test` ディレクトリに保存する。外部ファイルとして存在するデータセットを読み込む方法は、ファイル形式に応じて複数あるが、以下ではタブ区切りのテキスト形式を前提として説明する。外部ファイルの読み込み方法については、たとえば岡田編 (2004) を参照されたい。

外部ファイルとして作成するデータセットの構造は、1行に1つの選択肢のデータを保存する形式である。一般に調査結果のデータセットは、1行に回答者1名の回答結果が保存されるが、それとは異なるデータセット構造である。Greene (2000) や土木学会土木計画学研究委員会 (1995) などと同様なデータセット構造が紹介されているので必要な方は参照されたい。

Fig.9に、本事例のデータセットの一部を示す（全体は付録を参照）。列Aは回答者と質問の各組み合わせを区別（層化）するための変数STR，列Bは回答結果を表す変数RES，列Cは選択肢固有定数である変数ASC，列Dは1トン当たり堆肥価格を表す変数PRICE，列Eは堆肥の運搬・散布サービスの有無を表す変数TRANS，列Fは土壌診断・施肥設計サービスの有無を表す変数SOILである。

本事例では、1つの質問で提示される選択肢が3つであることから、3行で1問の情報となっている。データセットは上から回答者番号が1の人から10の人まで順に入力されている。さらに、回答者ごとに選択型コンジョイント分析の質問の回答結果と各選択肢で提示された水準が、問1から問8の順で入力されている。本事例では質問は全8問であり、1問当たりの選択肢数は3であることから、回答者1名分のデータは24行 (=8×3) にわたる。たとえば、行2から行25までの24行分には回答者1の回答結果が保存されており、行2が問1の堆肥A，

行3が問1の堆肥B，行4が問1の「どちらも買わない」の各選択肢で提示された水準（列Dの変数PRICE，列Eの変数TRANS，列Fの変数SOIL）と、それら3つの選択肢のうちいずれの選択肢が選ばれたかという情報（列Bの変数RES）等が保存されている。以下、各変数について詳しく説明する。

変数STRは、回答者と質問の組み合わせを識別（区別）するための変数である。各組み合わせを区別できればどのような値でもよい。ここでは、以下の式で変数STRを作成している。

$$STR = 100 \times \text{回答者番号} + \text{質問番号}$$

したがって、回答者1の問1では、

$$STR = 100 \times 1 + 1 = 101$$

となるし、回答者10の問8では、

$$STR = 100 \times 10 + 8 = 1008$$

となる。

変数RES（列B）は、いずれの選択肢が選ばれたかという選択結果を表す変数であり、選択された選択肢では「1」、それ以外の選択肢では「0」となるダミー変数形式である。たとえば、回答者1の問1に該当する3行（行2～行4）の変数RESをみると、3行目（行4）で「1」、残りの2行（行2と行3）では「0」の値となっており、回答者1は問1では3つめの選択肢、すなわち「どちらも買わない」を選んだことが確認できる。

変数ASC（列C）は選択肢固有定数であり、「どちらも買わない」選択肢のときを基準（0）として、堆肥Aと堆肥Bのときに「1」の値をとるように設定している。基本的な段階では、変数ASCは「どちらも買わない」選択肢では「0」、それ以外の具体的な選択肢ではすべて「1」に設定すると理解しておけばよい。

変数PRICE（列D）、変数TRANS（列E）、変数SOIL（列F）については、当該選択肢で提示された金額と各サービス内容（あり、なし）をダミー変数化した値が入力されているが、「どちらも買わない」選択肢に該当する行については、すべて「0」としている。今回使用する関数 `cllogit()` では「どちらも買わない」といった「非選択」選択肢を含む質問形式に対応していないので、データセット側で非選択選択肢の確定効用を「0」になるように構築する必要がある。基本的な段階では、「どちらも買わない」選択肢に対応する行の属性変数には「0」を入力すると理解しておけばよい。

### b データセットのRへの読み込み

データセットのRへの読み込み等を簡単にするため、作業ディレクトリをデータセットが保存されているデ

回答者番号は1

問1

回答者1は問1では「どちらも買わない」を選択

	A	B	C	D	E	F
1	STR	RES	ASC	PRICE	TRANS	SOIL
2	101	0	1	4000	1	1
3	101	0	1	2000	0	1
4	101	1	0	0	0	0
5	102	0	1	1000	0	0
6	102	1	1	1000	1	1
7	102	0	0	0	0	0
8	103	0	1	2000	0	1
9	103	0	1	4000	0	0
10	103	1	0	0	0	0
11	104	0	1	3000	0	1
12	104	0	1	1000	0	0
13	104	1	0	0	0	0
14	105	0	1	3000	1	0
15	105	0	1	3000	0	1
16	105	1	0	0	0	0
17	106	1	1	1000	1	1
18	106	0	1	4000	1	1
19	106	0	0	0	0	0
20	107	0	1	2000	1	0
21	107	0	1	3000	1	0
22	107	1	0	0	0	0
23	108	0	1	4000	0	0
24	108	0	1	2000	1	0
25	108	1	0	0	0	0
26	201	0	1	4000	1	1
27	201	0	1	2000	0	1

堆肥 A の各水準

堆肥 B の各水準

「どちらも買わない」

回答者1の問1から問8

131	604	0	1	3000	0	1
132	604	0	1	1000	0	0
133	604	1	0	0	0	0
134	605	0	1	3000	1	0
135	605	0	1	3000	0	1
136	605	1	0	0	0	0

回答者6の問4と問5

236	1007	0	1	2000	1	0
237	1007	0	1	3000	1	0
238	1007	1	0	0	0	0
239	1008	0	1	4000	0	0
240	1008	0	1	2000	1	0
241	1008	1	0	0	0	0

回答者10の問7と問8

Fig.9 関数clogit()で分析するためのデータセット(一部)  
A data set to demonstrate the function clogit()

ィレクトリーに変更する。作業ディレクトリーの変更には、関数setwd()を用いる。たとえば、作業ディレクトリーをcドライブ直下にあるtestディレクトリーに変更するためには、

```
setwd("c:/test")
```

と入力、実行すればよい。Fig.10では、その上で現在の作業ディレクトリーを表示する関数getwd()を実行して、作業ディレクトリーが実際にtestディレクトリーに変更されているか確認している。

データセットが文字の区切りにタブを使用したテキスト

```
> setwd("c:/test")
> getwd()
[1] "c:/test"
```

Fig.10 作業ディレクトリーの変更と確認  
The functions setwd() and getwd()

トデータであるため、データセットのRへの読み込みはbaseパッケージに含まれる関数read.delim()を用いて、以下のように入力、実行する。

```
data1<-read.delim("data1.txt")
```

```

> data1<-read.delim("data1.txt")
> data1
  STR RES ASC PRICE TRANS SOIL
1  101  0  1  4000    1    1
2  101  0  1  2000    0    1
3  101  1  0    0     0    0
4  102  0  1  1000    0    0
5  102  1  1  1000    1    1
6  102  0  0    0     0    0
7  103  0  1  2000    0    1
8  103  0  1  4000    0    0
9  103  1  0    0     0    0

```

Fig.11 データセットのRへの読み込み  
Importing data from an external file into R with  
the function read.delim()

ここでは、作業ディレクトリーにある「data1.txt」というファイルをRに読み込み、R内部では「data1」という名前と呼ぶように命令している。data1は、データフレーム形式となる。問題なくデータフレームdata1がR内部に構築できたか確認するためには、データセットの名前、

```
data1
```

を入力、実行すれば、データセットが表示される (Fig.11)。

#### c clogitの実行と結果の出力

Rではsurvivalパッケージに含まれる関数clogit()により、条件付きロジット・モデルを実行できる。関数clogit()の書式は、次の通りである。

```
clogit(被説明変数~説明変数+strata(層別変数), data=
データセット名)
```

今回のモデルでは、被説明変数は変数RES、説明変数は変数ASC、変数PRICE、変数TRANS、および変数

SOILが該当する。今回の事例のように説明変数が複数ある場合には、それぞれの説明変数の間に「+」を加えて入力する。層別変数とは、1つの質問で提示された選択肢集合を1つの組としてRが認識できるようにするための変数である。たとえば、今回のケースでは変数STRが層別変数に該当する。1つの質問で提示される選択肢の数は、「どちらも買わない」を含めて3つであることから、層別変数は3行ごとに異なる値が入力されている (Fig.9とFig.11)。もしも、1つの質問で提示される選択肢数が5であれば、層別変数には5行ごとに異なる値が入力される。最後の「data=」で分析するデータセットを指定する。今回のデータセット名はdata1である。

以上を踏まえて、計測結果をオブジェクトclogout1に保存するとすれば、関数library()によりsurvivalパッケージを呼び出した上で、本事例では次のようにコマンドを入力、実行することになる (Fig.12)。

```
clogout1<-clogit(RES~ASC+PRICE+TRANS+
SOIL+strata(STR),data=data1)
```

そしてオブジェクト名、

```
clogout1
```

を入力、実行すれば、計測結果が表示される。

#### d 計測結果の読み方

計測結果のうち注目するものは、各変数の係数推定値を表しているcoefの列、各係数推定値の有意確率を示すpの列、および尤度比検定統計量およびその有意確率を示すLikelihood ratio testの行である。

各変数の係数推定値の符号条件をみることで、各変数がそれぞれの選択肢の効用に与える影響がわかる。今回の各変数の定義から、係数推定値がプラスである変数は、その変数の値が大きくなるほど、その変数が設定されている選択肢の効用が大きくなることを意味する。逆に係数推定値がマイナスである変数は、その変数の値が大き

```

> clogout1<-clogit(RES~ASC+PRICE+TRANS+SOIL+strata(STR), data=data1)
> clogout1
Call:
clogit(RES ~ ASC + PRICE + TRANS + SOIL + strata(STR), data = data1)

      coef exp(coef) se(coef)      z      p
ASC    3.56470    35.329  1.15582  3.08 0.00200
PRICE -0.00351     0.997  0.00101 -3.48 0.00051
TRANS  3.17271    23.872  1.31931  2.40 0.01600
SOIL   3.54930    34.789  1.33246  2.66 0.00770

Likelihood ratio test=97.8 on 4 df, p=0 n= 240

```

Fig.12 関数clogit()による計測結果  
Results of the conditional logit model with the function clogit()



くなるほど、その変数が設定されている選択肢の効用が小さくなる。本結果についてみると、プラスの係数推定値を持つのは変数TRANSと変数SOIL、マイナスの係数推定値を持つのは変数PRICEである。運搬・散布サービスや土壌診断・施肥設計サービスが付加された堆肥を購入する方が、それらが付加されない堆肥を購入することよりも農家の効用は高くなる。一方、価格の高い堆肥であるほど、それを購入することで得られる農家の効用は低くなる。サービスが付加されて、価格の安い堆肥ほど農家にとって望ましいことがわかる。なお、選択肢固有定数(変数ASC)の符号については、基本的には気にする必要はない。

有意確率 $p$ とは、検定統計量の外側の確率のことである。その値が有意水準よりも小さければ、帰無仮説は棄却されると判断できる。たとえば、個々の係数推定値が0と有意差が認められるか検定する際の有意水準を10%と設定すれば、 $p$ の値が0.10より小さい係数推定値は0と統計的に有意差が認められることになる。今回の例(Fig.12)では、いずれの係数推定値の $p$ の値も0.10よりは小さいことから、少なくとも有意水準10%を基準とすれば、いずれも0と有意差が認められる。

尤度比検定(Likelihood ratio test)とは、「すべての係数推定値が0である」を帰無仮説とした統計的検定であり、尤度比検定統計量が自由度 $k$ (=係数推定値の個数)のカイ<sup>2</sup>乗分布に従うことを利用して行う。Likelihood ratio testと書かれた行にある $p$ の値が、この検定に対する有意確率である。今回の計測例では $p$ の値が0であり、高い確率で帰無仮説が棄却されることがわかる。

なお、標準では最大対数尤度が出力されないため、必要な場合は下記のように結果が保存されているオブジェ

クト名に続いて「\$loglik」を加えて、実行すれば初期対数尤度と合わせて出力される。

clogout1\$loglik

今回の事例では、初期対数尤度は-87.88898、最大対数尤度が-38.98682が出力される。

2 説明変数に個人・世帯特性が含まれるケース

a データセットの作成とRへの読み込み

説明変数に個人・世帯特性変数(以下、個人特性変数)が含まれるケースでも、データセットの基本構造は説明変数が属性変数のみのケースと同じである(Fig.13)。異なる点は、個人特性変数に該当する列が追加されることである。

Fig.13に示す例では、回答農家の経営耕地面積(列Gの変数SIZEで、単位はヘクタール)を個人特性変数として設定している。条件付きロジット・モデルで個人特性変数を分析に取り入れる1つの方法は、属性変数(変数ASCを含む)との交差項、つまり個人特性変数と属性変数の積としてモデルに取り入れるというものである。ここでは、堆肥の運搬・散布サービス(変数TRANS)と土壌診断・施肥設計サービス(変数SOIL)に対する評価(各変数の係数推定値)が経営耕地面積(変数SIZE)により変化すると仮定し、計測結果からその仮定を統計的に検証するとしてしよう。そこで、変数TRANSと変数SOILそれぞれと変数SIZEとの積を表す新しい変数をデータセットの中に作成する。

はじめに、変数STRから変数SIZEまでが保存されたデータセットが、ファイル名「data2.txt」とするタブ区切

	A	B	C	D	E	F	G
1	STR	RES	ASC	PRICE	TRANS	SOIL	SIZE
2	101	1	1	4000	1	1	1
3	101	0	1	2000	0	1	1
4	101	0	0	0	0	0	1
5	102	0	1	1000	0	0	1
6	102	1	1	1000	1	1	1

Fig.13 説明変数に個人・世帯特性変数が含まれるケースのデータセットの例

A data set including the respondents' characteristic

```

> data2<-read.delim("data2.txt")
> data2$TRANS_SIZE<-data2$TRANS*data2$SIZE
> data2$SOIL_SIZE<-data2$SOIL*data2$SIZE
> data2
  STR RES ASC PRICE TRANS SOIL SIZE TRANS_SIZE SOIL_SIZE
1  101  1  1  4000    1    1  1.0         1.0         1.0
2  101  0  1  2000    0    1  1.0         0.0         1.0
3  101  0  0    0     0    0  1.0         0.0         0.0
4  102  0  1  1000    0    0  1.0         0.0         0.0
    
```

Fig.14 個人特性変数を含むデータセットのRへの読み込みと確認(一部)

Importing a data set, including the respondents' characteristic from an external file into R and creating new variables



```

> clogout2<-clogit(RES~ASC+PRICE+TRANS+SOIL+TRANS_SIZE+SOIL_SIZE+strata(STR), data=data2)
> clogout2
Call:
clogit(RES ~ ASC + PRICE + TRANS + SOIL + TRANS_SIZE + SOIL_SIZE +
      strata(STR), data = data2)

            coef exp(coef) se(coef)      z      p
ASC          1.40645    4.081  1.51628  0.928 0.3500
PRICE        -0.00302    0.997  0.00109 -2.763 0.0057
TRANS         6.39502   598.856  2.54706  2.511 0.0120
SOIL          5.86668   353.076  2.68957  2.181 0.0290
TRANS_SIZE   1.02067    2.775  0.58518  1.744 0.0810
SOIL_SIZE    1.10244    3.012  0.56262  1.959 0.0500

Likelihood ratio test=135 on 6 df, p=0 n= 240

```

Fig.15 条件付きロジットモデルによる計測結果（個人特性変数を含むケース）

Results of the conditional logit model with the function `clogit()` (the case of a data set including the respondents' characteristic)

りのテキスト形式で「`c/test`」に保存されているとして、関数`read.delim()`を使ってRへ読み込む（Fig.14の1行目）。手順は、上述の説明変数が属性変数のみのケースと同様である。

ついで、変数`TRANS`と変数`SOIL`をそれぞれ変数`SIZE`と掛け合わせた新しい変数（変数`TRANS_SIZE`と変数`SOIL_SIZE`）をデータセットに作成する。これらの変数をRで作成するためには、次の2つのコマンドをそれぞれ入力し、実行する（Fig.14の2から3行目）。

```
data2$TRANS_SIZE<-data2$TRANS*data2$SIZE
```

```
data2$SOIL_SIZE<-data2$SOIL*data2$SIZE
```

ここで各変数名の前に「`data2$`」が付加されているのは、それぞれの変数が`data2`というデータセット（データフレーム）に保存されていることを示す。あるいは、そのデータセットに含まれる各変数を指定しているといってもよい\*9。

最後に、データセット名`data2`を入力、実行すれば、Rに読み込まれたデータセットと追加作成された変数`TRANS_SIZE`と変数`SOIL_SIZE`が確認できる（Fig.14の下方）。なお、追加された変数と選択結果を表す変数`RES`以外の変数については、説明変数が属性変数のみのケース（`data1`）と同じである。変数`RES`の値が異なるのは、新たに個人特性変数を含めたシミュレーションを行って生成したためである。

#### b clogitの実行と結果の出力

関数`clogit()`についても、基本的には説明変数が属性変数のみのケースと同じである。異なる部分は、説明変数として個人特性変数と属性変数との交差項を表す変数を含める点である。今回のケースでは、Fig.15に示すように、関数`clogit()`の説明変数を記載する部分で変数`SOIL`に続いて変数`TRANS_SIZE`と変数`SOIL_SIZE`の2つの変数を加える。また、データセット名は対応するデータフレーム名の`data2`とする。

計測結果をオブジェクト`clogout2`に保存するとして、実行した結果をFig.15の下方に示す。変数`TRANS_SIZE`と変数`SOIL_SIZE`のどちらの係数推定値（列`coef`）も符号はプラスである。また、それらの係数推定値のp値はどちらも0.10より小さく、有意水準10%であれば0と有意差が認められる。これらの結果より、経営耕地面積（`SIZE`）が大きい回答農家ほど、堆肥の運搬・散布サービスや土壌診断・施肥設計サービスに対して高い評価を下し、それらのサービスが付いている堆肥を選択する傾向がより強くなることがわかる。

## 結 言

本稿では、データ解析環境Rを用いた選択型コンジョイント分析の実施手順について具体的に紹介した。使用するコマンドは少数に限定されているとはいえ、コマンドベースでのソフトウェアの操作に抵抗感のある初心者も少なからず存在しよう。今後の課題としては、Rで選択型コンジョイント分析がより簡単に実行できるようなGUIベースのインターフェースの作成がある。

\*9 関数`attach()`を使えば、もう少し簡略化することもできる。また、関数`clogit()`でモデル式を指定する際に、記号：（コロン）を使うことで交差項を事前に作成しない方法もある。

## 補 論

## 1 実験回数について

本文の解説では、16回の完全実施要因計画を8回の部分実施要因計画に縮減したが、9回や7回といったように異なる回数の部分実施要因計画とすることも可能である。ただし、実験回数を無制限に小さくすることはできない。たとえば、今回のケースで5回に設定して関数 `optFederov()` を実行すると、「以下にエラー `optFederov(～., ffd, 5) : nTrials must be greater than or equal to the number of columns in expanded X`」が表示され、部分実施要因計画を作成できないことがわかる。

また、実験回数によっては各要因が直交しない、つまり要因間に非ゼロの相関が生じることもある。たとえば、以下のように10回、6回の2種類の部分実施要因計画を作成し、8回の部分実施要因計画とあわせて、それぞれの計画で要因間の相関係数を求めてみよう。

```
des10<-optFederov(～.,ffd,10)
des6<-optFederov(～.,ffd,6)
```

Fig.16に示すように `stats` パッケージ (Rと同時に読み込み済み) に含まれる関数 `cor()` を用いて要因間の相関係数を求めると、8回の部分実施要因計画では相関係数はいずれもゼロであるが、10回と6回の部分実施要因計画では要因間の一部あるいはすべてに非ゼロの相関が生じている。8回のケースにおいて各相関係数がいずれもゼロであるのは、ここで取り上げた完全実施要因計画の厳密な部分実施要因計画になっている (対応する直交配列表が存在する) ためである。

一般に、完全実施要因計画からの実験回数の削減は変数間の相関をもたらす。離散選択モデルを前提とした選択型コンジョイント分析では、直交性以外の条件も含めて選択肢集合の最適性を評価すべきという議論もあるため、程度の問題を無視すれば属性間に相関が生じること自体は大きな問題ではない<sup>\*10</sup>。

それでは、今回の全選択肢集合の作成方法を前提として、実験回数 (質問回数) は何回が適当なのかという疑問が生じると思うが、全ての条件にとって最適であるという解は存在しないだろう。たとえば、選択肢を構成する属性数が多く、回答に際して回答者に負荷を強いるケースでは、無回答やいい加減な回答が増えることを避け

\*10 たとえば、選択型コンジョイント分析を用いた近年の研究では、今回用いたものと類似のアルゴリズム (離散選択モデルを前提としたD効率性の最適化) によって直交配列表や完全実施要因計画から実験回数を削減しているが、直交性以外の条件も勘案して最適化しているため、変数間には非ゼロの相関を持っている。関心のある方はHuber and Zwerina(1996), Zwerina et al.(1996)を参照されたい。

```
> des10<-optFederov(～., ffd, 10)
> des6<-optFederov(～., ffd, 6)
> des10$design
  PRICE TRANS SOIL
2      2      1      1
3      3      1      1
4      4      1      1
5      1      2      1
6      2      2      1
7      3      2      1
9      1      1      2
10     2      1      2
15     3      2      2
16     4      2      2
> cor(des10$design)
      PRICE  TRANS  SOIL
PRICE 1.00000 0.09759    0
TRANS 0.09759 1.00000    0
SOIL  0.00000 0.00000    1
> des$design
  PRICE TRANS SOIL
1      1      1      1
4      4      1      1
6      2      2      1
7      3      2      1
10     2      1      2
11     3      1      2
13     1      2      2
16     4      2      2
> cor(des$design)
      PRICE  TRANS  SOIL
PRICE      1      0      0
TRANS      0      1      0
SOIL       0      0      1
> des6$design
  PRICE TRANS SOIL
2      2      1      1
3      3      1      1
5      1      2      1
9      1      1      2
15     3      2      2
16     4      2      2
> cor(des6$design)
      PRICE  TRANS  SOIL
PRICE 1.0000000 0.3015113 0.3015113
TRANS 0.3015113 1.0000000 0.3333333
SOIL  0.3015113 0.3333333 1.0000000
```

Fig.16 実施回数と要因間の相関係数  
Relationship between the number of trials and  
the correlation among variables

るために、回答者1人当たり質問回数を減らす必要がある。1人当たり質問回数を減らすためには、全選択肢集合を構成する選択肢集合の数、すなわち質問 (実験回数) の総数を減らすか、選択肢集合をいくつか分割し、調査票によって提示する質問の内容 (属性水準の組み合わせ) を変化させて、1人当たり質問回数を減らす方策がある。たとえば、合崎 (2005) に掲載されている4事例では、いずれもの方法で回答者1人当たり質問回数を減らしている。

前者の方策を用いるのであれば、属性間の相関をある程度認めた上で、関数`optFederov()`を実施するときの目標とする実験回数を少なくする。この場合、どの程度の相関を認めるかという疑問が生じるかもしれないが、モンテカルロ・シミュレーションを実施する必要があるなどはじめて選択型コンジョイント分析に取り組むような人では定量的に評価するのが難しいかもしれない。そのような人でも実施可能な検討方法の1つは、予備調査を実施して非現実的な計測結果にならないか確認するという方法である。

後者の方策を選択するのであれば、属性間の相関が生じないような組み合わせを採用することもできるかもしれない。ただし、こちらの場合、複数のバージョンの調査票が作成されるため、バージョンによって有効回答率に相違が生じ、そのことが属性間の相関を引き起こす可能性がある（たとえば、合崎2006）。どちらの方策を採用するかは、人それぞれの考え方に依存するだろう。

以上を踏まえれば、選択型コンジョイント分析を実施しようとする際には、どのような考え方に基づいて選択肢集合を作成するのか自ら決定し、作成した全選択肢集合を論文に掲載したり、問い合わせがあれば提供する旨を記載したりすること、と整理できよう。採用した選択肢集合の作成方針が全ての人に受け入れられることはないだろう。したがって、そこには議論の余地が生まれる。しかし、選択肢集合の作成方法を明記するだけでは、属性間にどの程度の相関が生じたのか、どのような組み合わせの実験になったのかは不明確である。全選択肢集合をオープンにすれば、どのような条件の下で実施したのかを他者が正確に確認でき、適切な検討や議論を展開することが可能となる。

## 2 計測結果の活用について

選択型コンジョイント分析を活用する場合、係数推定値を得ることが目的ではなく、係数推定値を使ってさまざまな情報を引き出すことが目的となることが多い。以下では、限界支払意思額と選択確率の予測の2つを紹介する。これら以外にも、複数の製品を設定して、それらの選択確率が等しくなる（効用が無差別になる）価格水準を求めたり、個人特性による各種評価額や予測値の違いを比較したりすることができる。計測結果を使ってどのようなことが可能であるかは、さまざまな既存文献を参照されたい。

### a 限界支払意思額（属性の価格評価）

非価格属性変数の限界支払意思額は、非価格属性変数が1単位変化したときの評価額を表す。非価格属性が、回答者からみてどの程度の経済価値を持つのかを貨幣単位で表したものである。線形モデルであれば、非価格属性変数の係数推定値を価格属性変数の係数推定値で除して「 $\cdot$ 」を乗じることで得られる。

今回の例であれば、非価格属性変数として堆肥の運

搬・散布サービスや土壌診断・施肥設計サービスが設定されていることから、それらサービスに対する農家の評価額が得られる。Fig.12より価格（変数PRICE）の係数推定値は「-0.00351」、堆肥の運搬・散布サービス属性変数（変数TRANS）の係数推定値は「3.17271」、土壌診断・施肥設計サービス属性変数（変数SOIL）の係数推定値は「3.54930」であることから、

運搬・散布サービスの農家評価額

$$= -1 \times 3.17271 \div -0.00351 \\ 904 \text{円}$$

土壌診断・施肥設計サービスの農家評価額

$$= -1 \times 3.54930 \div -0.00351 \\ 1,011 \text{円}$$

と求められる。

### b 選択確率の予測

計測結果を用いることで特定の選択肢の選択確率を求めることができ、選択確率が予測できればさらに需要量を予測することもできる（たとえば、合崎（2002, 2005）を参照）。たとえば、堆肥の運搬・散布サービスと土壌診断・施肥設計サービスの付いた堆肥の選択確率（購入されると予想される確率）を予測してみよう。堆肥市場には、この堆肥1種類のみが販売されると仮定して、稲作農家の取り得る行動は堆肥を「購入する」か「購入しない」かの2つであるとする。価格を変数 $x$ で表すとすれば、計測結果から堆肥の購入確率 $p_{pb}$ は、次のように定式化できる。

$$p_{pb} = \exp(3.5647 - 0.00351 \cdot x + 3.17271 + 3.5493) \div \\ (\exp(3.5647 - 0.00351 \cdot x + 3.17271 + 3.5493) + \exp(0))$$

分子は、堆肥を購入することから得られる確定効用を指数変換（ $\exp$ ）したもので、分母は購入することから得られる確定効用を指数変換したものと、購入しないことから得られる確定効用を指数変換したものの（ $\exp(0)=1$ ）の和である。評価対象の堆肥には2つのサービスが付加されているため、堆肥購入に対応する確定効用には選択肢固有定数項（3.5647）に加えて、それら2つの係数推定値（3.17271と3.5493）が加わっている。この式を適当な $x$ の範囲で計算することで、堆肥の購入確率 $p_{pb}$ を予測できる。ここでは、堆肥の1トン当たり価格が1,000円から4,000円の範囲での堆肥の購入確率と、さらに地域全体での堆肥需要量を $R$ を使って予測（作図）してみよう。

$R$ で散布図を描くためにはいくつかの関数があるが、ここでは関数`plot()`を使う。通常、確率を図示する際には、横軸に変化させる要因（今回であれば価格を表す変数 $x$ ）、縦軸が確率（今回であれば購入確率 $p_{pb}$ ）と



```

> prb<-function(x) exp(3.56470-0.00351*x+3.17271+3.54930)/(exp(3.56470-0.00351*x+3.17271+3.54930)+exp(0))
> plot(prb(1000:4000), 1000:4000, type="l", xlab="Choice Probability", ylab="PRICE")
> plot(100*prb(1000:4000)*3*20, 1000:4000, type="l", xlab="Demand", ylab="PRICE")

```

Fig.17 関数plot()による堆肥の購入確率曲線と需要曲線の作図

Drawing a choice probability curve and a demand curve for compost with the function plot()

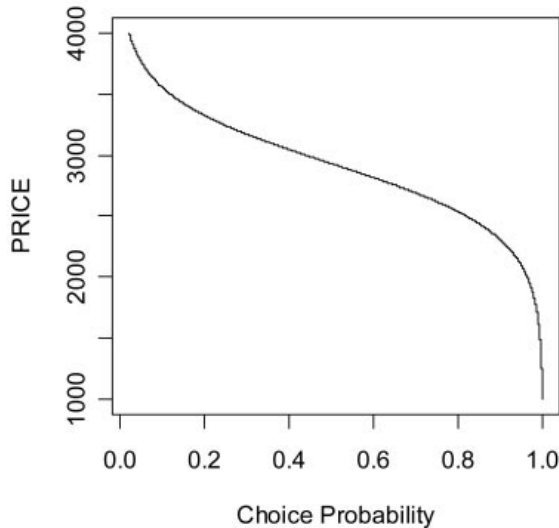


Fig.18 堆肥の購入確率曲線

A choice probability curve for compost

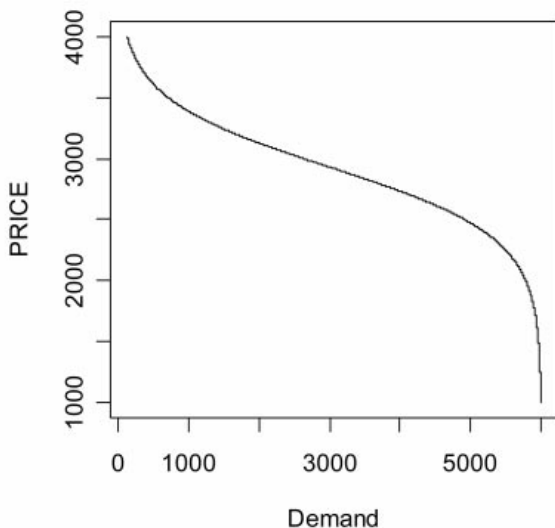


Fig.19 堆肥の需要曲線

A demand curve for compost

なる。しかし、堆肥の需要量も予測することから、ここでは縦軸を価格、横軸を確率として描く。経済学では縦軸を価格、横軸を需要量として需要曲線（関数）を描くことが一般的のためである。

Fig.17の1行目に示すように、はじめに堆肥の選択確率 $prb$ を関数として定義する。Rで任意の関数を設定するときは関数`function()`を用いるが、使い方の詳細は船尾(2005)などを参照されたい。つぎに2行目の関数`plot()`

を使って散布図を描く。関数`plot()`の括弧内の第1要素(`prb(1000:4000)`)が横軸を、第2要素(`1000:4000`)が縦軸を、第3要素(`type="l"`)が線を使って描写する(折れ線図)ことを、第4要素(`xlab="Choice Probability"`)が横軸の名前を、第5要素(`ylab="PRICE"`)が縦軸の名前を、それぞれ示している。これらを実行することで、Fig.18に示す堆肥の購入確率曲線が描ける。

ついで、堆肥の需要曲線を描いてみよう(Fig.17の3行目およびFig.19)。ここでは仮に、本事例のデータが農家100戸(潜在的な堆肥購入主体)の地域から無作為に抽出されたものであるとしよう。1戸当たり平均経営耕地面積が3ヘクタール、1ヘクタール当たり平均堆肥投入量が20トンであるとする(いずれも仮想値)。堆肥の購入確率を堆肥の購入農家率と読み替えて、この地域全体の堆肥需要量 $D$ を以下のように定式化できる。

$$D = 100 \times prb \times 3 \times 20$$

堆肥を購入すると予想される農家戸数は、地域全体の農家戸数(100戸)に堆肥の購入農家率 $prb$ を乗じることで得られる。予想された堆肥購入農家戸数に平均経営耕地面積(3ヘクタール)を乗じることで、地域全体での堆肥が投入される経営耕地面積が求まる。そして、その面積に1ヘクタール当たり平均堆肥投入量(20トン)を乗じれば、地域全体としての堆肥需要量が予測できるという考え方である。この考え方に基づいた関数`plot()`のコマンドがFig.17の3行目、その結果がFig.19である。なお、例示を簡単にするため経営耕地面積等の農家特性による堆肥購入確率の違いは無視している点には注意されたい。

## 参考文献

- 1) 合崎英男(2004):直交配列表を用いたプロフィール作成の手順,農業工学研究所技報,200,21-32
- 2) 合崎英男(2005):農業・農村の計画評価-表明選好法による接近-,農林統計協会
- 3) 合崎英男(2006):生ごみ分別・回収活動における他世帯協力率と自己の協力意向,農業情報研究,15(1),1-14
- 4) Bennett, J. and R. Blamey (2001): The Choice Modelling Approach to Environmental Valuation, Edward Elgar
- 5) Champ, P. A., K. J. Boyle and T. C. Brown (2003): A



- Primer on Nonmarket Valuation, Kluwer Academic Publishers
- 6) 土木学会土木計画学研究委員会 (1995): 非集計行動モデルの理論と実際, 土木学会
- 7) 船尾暢男 (2005): The R Tips データ解析環境Rの基本技・グラフィック活用集, 九天社
- 8) 船尾暢男・高浪洋平 (2005): データ解析環境「R」, 工学社
- 9) Greene, W. H. 著, 斯波恒正・中妻照雄・浅井 学訳 (2000): グリーン計量経済学, エコノミスト社
- 10) Hensher, D. A., J. M. Rose and W. H. Greene (2005): Applied Choice Analysis, Cambridge University Press
- 11) Huber, J. and K. Zwerina (1996): The Importance of Utility Balance in Efficient Choice Design, Journal of Marketing Research, 33, 307-317
- 12) 栗山浩一 (2000): SPによる交通需要評価とその統計的効率性, 運輸政策研究, 3(2), 2-8
- 13) 栗山浩一・庄子 康編著 (2005): 環境と観光の経済評価 - 国立公園の維持と管理 -, 勁草書房
- 14) Louviere, J. J., D. A. Hensher, and J. D. Swait (2000): Stated Choice Methods, Cambridge University Press
- 15) Lunley, T. (2006): clogit. S original by T. Theureau, ported by T. Lunley: The survival package version 2.29
- 16) 岡田昌史編 (2004): The R Book データ解析環境Rの活用事例集, 九天社
- 17) R Development Core Team (2006): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <http://www.R-project.org>
- 18) 柘植隆宏・栗山浩一・庄子 康 (2005): 環境評価手法としてのコンジョイント分析, 栗山浩一・庄子康編著: 環境と観光の経済評価 - 国立公園の維持と管理 -, 勁草書房, 63-94
- 19) Wheeler, R. E. (2004): Comments on algorithmic design
- 20) Wheeler, B. (2006): AlgDesign: AlgDesign. R package version 1.0-7. <http://www.bobwheeler.com/stat>
- 21) Zwerina, K., J. Huber, and W. F. Kuhfeld (1996): A General Method for Constructing Efficient Choice Designs, SAS Technical Support Documents, TS-694E

# Introduction to Choice Experiments Using R

AIZAKI Hideo and NISHIMURA Kazushi

## Summary

This paper is an introduction to choice experiments (choice-based conjoint analysis) using R, which is a language and environment for statistical computing. Numerous statistical methods that are provided as packages can be added to R. This document comprises two parts; The first is an explanation of the process of creating a choice experiments survey (choice sets) with the functions `gen.factorial()` and `optFederov()` included in the `AlgDesign` package, and the second is an explanation of the process of analyzing data gathered by the survey with the function `clogit()` included in the `survival` package. In addition, we have provided information regarding the estimating of marginal willingness to pay for non-price attributes and the forecasting of choice probability of an alternative with R. Since R is a free software, many readers of this document may find it easy to apply the choice experiments to their research projects.

**Keywords** :R, conjoint analysis, choice experiments, design of experiments, choice sets, discrete choice model, conditional logit model

## 付録1 説明変数が属性変数のみのデータセット

STR	RES	ASC	PRICE	TRANS	SOIL
101	0	1	4000	1	1
101	0	1	2000	0	1
101	1	0	0	0	0
102	0	1	1000	0	0
102	1	1	1000	1	1
102	0	0	0	0	0
103	0	1	2000	0	1
103	0	1	4000	0	0
103	1	0	0	0	0
104	0	1	3000	0	1
104	0	1	1000	0	0
104	1	0	0	0	0
105	0	1	3000	1	0
105	0	1	3000	0	1
105	1	0	0	0	0
106	1	1	1000	1	1
106	0	1	4000	1	1
106	0	0	0	0	0
107	0	1	2000	1	0
107	0	1	3000	1	0
107	1	0	0	0	0
108	0	1	4000	0	0
108	0	1	2000	1	0
108	1	0	0	0	0
201	0	1	4000	1	1
201	0	1	2000	0	1
201	1	0	0	0	0
202	0	1	1000	0	0
202	1	1	1000	1	1
202	0	0	0	0	0
203	1	1	2000	0	1
203	0	1	4000	0	0
203	0	0	0	0	0
204	0	1	3000	0	1
204	1	1	1000	0	0
204	0	0	0	0	0
205	0	1	3000	1	0
205	0	1	3000	0	1
205	1	0	0	0	0
206	1	1	1000	1	1
206	0	1	4000	1	1
206	0	0	0	0	0
207	1	1	2000	1	0
207	0	1	3000	1	0
207	0	0	0	0	0
208	0	1	4000	0	0
208	0	1	2000	1	0
208	1	0	0	0	0
301	0	1	4000	1	1
301	0	1	2000	0	1
301	1	0	0	0	0
302	0	1	1000	0	0
302	1	1	1000	1	1
302	0	0	0	0	0
303	0	1	2000	0	1
303	0	1	4000	0	0
303	1	0	0	0	0
304	0	1	3000	0	1
304	0	1	1000	0	0
304	1	0	0	0	0
305	0	1	3000	1	0

305	0	1	3000	0	1
305	1	0	0	0	0
306	1	1	1000	1	1
306	0	1	4000	1	1
306	0	0	0	0	0
307	0	1	2000	1	0
307	0	1	3000	1	0
307	1	0	0	0	0
308	0	1	4000	0	0
308	1	1	2000	1	0
308	0	0	0	0	0
401	0	1	4000	1	1
401	1	1	2000	0	1
401	0	0	0	0	0
402	0	1	1000	0	0
402	1	1	1000	1	1
402	0	0	0	0	0
403	1	1	2000	0	1
403	0	1	4000	0	0
403	0	0	0	0	0
404	0	1	3000	0	1
404	0	1	1000	0	0
404	1	0	0	0	0
405	0	1	3000	1	0
405	0	1	3000	0	1
405	1	0	0	0	0
406	1	1	1000	1	1
406	0	1	4000	1	1
406	0	0	0	0	0
407	1	1	2000	1	0
407	0	1	3000	1	0
407	0	0	0	0	0
408	0	1	4000	0	0
408	1	1	2000	1	0
408	0	0	0	0	0
501	0	1	4000	1	1
501	1	1	2000	0	1
501	0	0	0	0	0
502	0	1	1000	0	0
502	1	1	1000	1	1
502	0	0	0	0	0
503	1	1	2000	0	1
503	0	1	4000	0	0
503	0	0	0	0	0
504	0	1	3000	0	1
504	1	1	1000	0	0
504	0	0	0	0	0
505	0	1	3000	1	0
505	0	1	3000	0	1
505	1	0	0	0	0
506	1	1	1000	1	1
506	0	1	4000	1	1
506	0	0	0	0	0
507	0	1	2000	1	0
507	0	1	3000	1	0
507	1	0	0	0	0
508	0	1	4000	0	0
508	0	1	2000	1	0
508	1	0	0	0	0
601	0	1	4000	1	1
601	0	1	2000	0	1
601	1	0	0	0	0

602	0	1	1000	0	0
602	1	1	1000	1	1
602	0	0	0	0	0
603	1	1	2000	0	1
603	0	1	4000	0	0
603	0	0	0	0	0
604	0	1	3000	0	1
604	0	1	1000	0	0
604	1	0	0	0	0
605	0	1	3000	1	0
605	0	1	3000	0	1
605	1	0	0	0	0
606	1	1	1000	1	1
606	0	1	4000	1	1
606	0	0	0	0	0
607	0	1	2000	1	0
607	0	1	3000	1	0
607	1	0	0	0	0
608	0	1	4000	0	0
608	0	1	2000	1	0
608	1	0	0	0	0
701	0	1	4000	1	1
701	1	1	2000	0	1
701	0	0	0	0	0
702	0	1	1000	0	0
702	1	1	1000	1	1
702	0	0	0	0	0
703	1	1	2000	0	1
703	0	1	4000	0	0
703	0	0	0	0	0
704	0	1	3000	0	1
704	1	1	1000	0	0
704	0	0	0	0	0
705	1	1	3000	1	0
705	0	1	3000	0	1
705	0	0	0	0	0
706	1	1	1000	1	1
706	0	1	4000	1	1
706	0	0	0	0	0
707	1	1	2000	1	0
707	0	1	3000	1	0
707	0	0	0	0	0
708	0	1	4000	0	0
708	0	1	2000	1	0
708	1	0	0	0	0
801	0	1	4000	1	1
801	0	1	2000	0	1
801	1	0	0	0	0
802	0	1	1000	0	0
802	1	1	1000	1	1
802	0	0	0	0	0
803	1	1	2000	0	1
803	0	1	4000	0	0
803	0	0	0	0	0
804	0	1	3000	0	1
804	0	1	1000	0	0
804	1	0	0	0	0
805	0	1	3000	1	0
805	0	1	3000	0	1
805	1	0	0	0	0
806	1	1	1000	1	1
806	0	1	4000	1	1

806	0	0	0	0	0
807	1	1	2000	1	0
807	0	1	3000	1	0
807	0	0	0	0	0
808	0	1	4000	0	0
808	1	1	2000	1	0
808	0	0	0	0	0
901	0	1	4000	1	1
901	1	1	2000	0	1
901	0	0	0	0	0
902	0	1	1000	0	0
902	1	1	1000	1	1
902	0	0	0	0	0
903	0	1	2000	0	1
903	0	1	4000	0	0
903	1	0	0	0	0
904	0	1	3000	0	1
904	1	1	1000	0	0
904	0	0	0	0	0
905	0	1	3000	1	0
905	0	1	3000	0	1
905	1	0	0	0	0
906	1	1	1000	1	1
906	0	1	4000	1	1
906	0	0	0	0	0
907	0	1	2000	1	0
907	0	1	3000	1	0
907	1	0	0	0	0
908	0	1	4000	0	0
908	1	1	2000	1	0
908	0	0	0	0	0
1001	0	1	4000	1	1
1001	1	1	2000	0	1
1001	0	0	0	0	0
1002	0	1	1000	0	0
1002	1	1	1000	1	1
1002	0	0	0	0	0
1003	0	1	2000	0	1
1003	0	1	4000	0	0
1003	1	0	0	0	0
1004	0	1	3000	0	1
1004	1	1	1000	0	0
1004	0	0	0	0	0
1005	0	1	3000	1	0
1005	0	1	3000	0	1
1005	1	0	0	0	0
1006	1	1	1000	1	1
1006	0	1	4000	1	1
1006	0	0	0	0	0
1007	0	1	2000	1	0
1007	0	1	3000	1	0
1007	1	0	0	0	0
1008	0	1	4000	0	0
1008	0	1	2000	1	0
1008	1	0	0	0	0



## 付録2 説明変数に個人特性変数を含むデータセット

STR	RES	ASC	PRICE	TRANS	SOIL	SIZE
101	1	1	4000	1	1	1
101	0	1	2000	0	1	1
101	0	0	0	0	0	1
102	0	1	1000	0	0	1
102	1	1	1000	1	1	1
102	0	0	0	0	0	1
103	1	1	2000	0	1	1
103	0	1	4000	0	0	1
103	0	0	0	0	0	1
104	0	1	3000	0	1	1
104	0	1	1000	0	0	1
104	1	0	0	0	0	1
105	0	1	3000	1	0	1
105	0	1	3000	0	1	1
105	1	0	0	0	0	1
106	1	1	1000	1	1	1
106	0	1	4000	1	1	1
106	0	0	0	0	0	1
107	1	1	2000	1	0	1
107	0	1	3000	1	0	1
107	0	0	0	0	0	1
108	0	1	4000	0	0	1
108	1	1	2000	1	0	1
108	0	0	0	0	0	1
201	1	1	4000	1	1	1.5
201	0	1	2000	0	1	1.5
201	0	0	0	0	0	1.5
202	0	1	1000	0	0	1.5
202	1	1	1000	1	1	1.5
202	0	0	0	0	0	1.5
203	1	1	2000	0	1	1.5
203	0	1	4000	0	0	1.5
203	0	0	0	0	0	1.5
204	1	1	3000	0	1	1.5
204	0	1	1000	0	0	1.5
204	0	0	0	0	0	1.5
205	0	1	3000	1	0	1.5
205	0	1	3000	0	1	1.5
205	1	0	0	0	0	1.5
206	1	1	1000	1	1	1.5
206	0	1	4000	1	1	1.5
206	0	0	0	0	0	1.5
207	1	1	2000	1	0	1.5
207	0	1	3000	1	0	1.5
207	0	0	0	0	0	1.5
208	0	1	4000	0	0	1.5
208	1	1	2000	1	0	1.5
208	0	0	0	0	0	1.5
301	1	1	4000	1	1	2
301	0	1	2000	0	1	2
301	0	0	0	0	0	2
302	0	1	1000	0	0	2
302	1	1	1000	1	1	2
302	0	0	0	0	0	2
303	1	1	2000	0	1	2
303	0	1	4000	0	0	2
303	0	0	0	0	0	2
304	1	1	3000	0	1	2
304	0	1	1000	0	0	2
304	0	0	0	0	0	2
305	0	1	3000	1	0	2

305	0	1	3000	0	1	2
305	1	0	0	0	0	2
306	1	1	1000	1	1	2
306	0	1	4000	1	1	2
306	0	0	0	0	0	2
307	1	1	2000	1	0	2
307	0	1	3000	1	0	2
307	0	0	0	0	0	2
308	0	1	4000	0	0	2
308	1	1	2000	1	0	2
308	0	0	0	0	0	2
401	1	1	4000	1	1	2.5
401	0	1	2000	0	1	2.5
401	0	0	0	0	0	2.5
402	0	1	1000	0	0	2.5
402	1	1	1000	1	1	2.5
402	0	0	0	0	0	2.5
403	1	1	2000	0	1	2.5
403	0	1	4000	0	0	2.5
403	0	0	0	0	0	2.5
404	1	1	3000	0	1	2.5
404	0	1	1000	0	0	2.5
404	0	0	0	0	0	2.5
405	1	1	3000	1	0	2.5
405	0	1	3000	0	1	2.5
405	0	0	0	0	0	2.5
406	1	1	1000	1	1	2.5
406	0	1	4000	1	1	2.5
406	0	0	0	0	0	2.5
407	1	1	2000	1	0	2.5
407	0	1	3000	1	0	2.5
407	0	0	0	0	0	2.5
408	0	1	4000	0	0	2.5
408	1	1	2000	1	0	2.5
408	0	0	0	0	0	2.5
501	1	1	4000	1	1	3
501	0	1	2000	0	1	3
501	0	0	0	0	0	3
502	0	1	1000	0	0	3
502	1	1	1000	1	1	3
502	0	0	0	0	0	3
503	1	1	2000	0	1	3
503	0	1	4000	0	0	3
503	0	0	0	0	0	3
504	1	1	3000	0	1	3
504	0	1	1000	0	0	3
504	0	0	0	0	0	3
505	0	1	3000	1	0	3
505	1	1	3000	0	1	3
505	0	0	0	0	0	3
506	1	1	1000	1	1	3
506	0	1	4000	1	1	3
506	0	0	0	0	0	3
507	1	1	2000	1	0	3
507	0	1	3000	1	0	3
507	0	0	0	0	0	3
508	0	1	4000	0	0	3
508	1	1	2000	1	0	3
508	0	0	0	0	0	3
601	1	1	4000	1	1	3.5
601	0	1	2000	0	1	3.5
601	0	0	0	0	0	3.5

602	0	1	1000	0	0	3.5
602	1	1	1000	1	1	3.5
602	0	0	0	0	0	3.5
603	1	1	2000	0	1	3.5
603	0	1	4000	0	0	3.5
603	0	0	0	0	0	3.5
604	1	1	3000	0	1	3.5
604	0	1	1000	0	0	3.5
604	0	0	0	0	0	3.5
605	1	1	3000	1	0	3.5
605	0	1	3000	0	1	3.5
605	0	0	0	0	0	3.5
606	1	1	1000	1	1	3.5
606	0	1	4000	1	1	3.5
606	0	0	0	0	0	3.5
607	1	1	2000	1	0	3.5
607	0	1	3000	1	0	3.5
607	0	0	0	0	0	3.5
608	0	1	4000	0	0	3.5
608	1	1	2000	1	0	3.5
608	0	0	0	0	0	3.5
701	1	1	4000	1	1	4
701	0	1	2000	0	1	4
701	0	0	0	0	0	4
702	0	1	1000	0	0	4
702	1	1	1000	1	1	4
702	0	0	0	0	0	4
703	1	1	2000	0	1	4
703	0	1	4000	0	0	4
703	0	0	0	0	0	4
704	0	1	3000	0	1	4
704	1	1	1000	0	0	4
704	0	0	0	0	0	4
705	1	1	3000	1	0	4
705	0	1	3000	0	1	4
705	0	0	0	0	0	4
706	1	1	1000	1	1	4
706	0	1	4000	1	1	4
706	0	0	0	0	0	4
707	1	1	2000	1	0	4
707	0	1	3000	1	0	4
707	0	0	0	0	0	4
708	0	1	4000	0	0	4
708	1	1	2000	1	0	4
708	0	0	0	0	0	4
801	1	1	4000	1	1	4.5
801	0	1	2000	0	1	4.5
801	0	0	0	0	0	4.5
802	0	1	1000	0	0	4.5
802	1	1	1000	1	1	4.5
802	0	0	0	0	0	4.5
803	1	1	2000	0	1	4.5
803	0	1	4000	0	0	4.5
803	0	0	0	0	0	4.5
804	1	1	3000	0	1	4.5
804	0	1	1000	0	0	4.5
804	0	0	0	0	0	4.5
805	0	1	3000	1	0	4.5
805	1	1	3000	0	1	4.5
805	0	0	0	0	0	4.5
806	1	1	1000	1	1	4.5
806	0	1	4000	1	1	4.5

806	0	0	0	0	0	4.5
807	1	1	2000	1	0	4.5
807	0	1	3000	1	0	4.5
807	0	0	0	0	0	4.5
808	0	1	4000	0	0	4.5
808	1	1	2000	1	0	4.5
808	0	0	0	0	0	4.5
901	1	1	4000	1	1	5
901	0	1	2000	0	1	5
901	0	0	0	0	0	5
902	0	1	1000	0	0	5
902	1	1	1000	1	1	5
902	0	0	0	0	0	5
903	1	1	2000	0	1	5
903	0	1	4000	0	0	5
903	0	0	0	0	0	5
904	1	1	3000	0	1	5
904	0	1	1000	0	0	5
904	0	0	0	0	0	5
905	1	1	3000	1	0	5
905	0	1	3000	0	1	5
905	0	0	0	0	0	5
906	1	1	1000	1	1	5
906	0	1	4000	1	1	5
906	0	0	0	0	0	5
907	1	1	2000	1	0	5
907	0	1	3000	1	0	5
907	0	0	0	0	0	5
908	0	1	4000	0	0	5
908	1	1	2000	1	0	5
908	0	0	0	0	0	5
1001	1	1	4000	1	1	5.5
1001	0	1	2000	0	1	5.5
1001	0	0	0	0	0	5.5
1002	0	1	1000	0	0	5.5
1002	1	1	1000	1	1	5.5
1002	0	0	0	0	0	5.5
1003	1	1	2000	0	1	5.5
1003	0	1	4000	0	0	5.5
1003	0	0	0	0	0	5.5
1004	1	1	3000	0	1	5.5
1004	0	1	1000	0	0	5.5
1004	0	0	0	0	0	5.5
1005	0	1	3000	1	0	5.5
1005	1	1	3000	0	1	5.5
1005	0	0	0	0	0	5.5
1006	1	1	1000	1	1	5.5
1006	0	1	4000	1	1	5.5
1006	0	0	0	0	0	5.5
1007	0	1	2000	1	0	5.5
1007	1	1	3000	1	0	5.5
1007	0	0	0	0	0	5.5
1008	0	1	4000	0	0	5.5
1008	1	1	2000	1	0	5.5
1008	0	0	0	0	0	5.5

## 付録3 AlgDesignパッケージによる選択肢集合の作成（含む補論1）で入力，実行したコマンド

```

library(AlgDesign)
ffd<-gen.factorial(c(4,2,2), varNames=c("PRICE","TRANS","SOIL"), factors="all")
ffd
runif(1)
des<-optFederov(~., ffd, 8)
des
cs1<-des$design
cs1
des10<-optFederov(~., ffd, 10)
des6<-optFederov(~., ffd, 6)
des10$design
cor(des10$design)
des$design
cor(des$design)
des6$design
cor(des6$design)
cs2<-cs1
cs2
cs1<-transform(cs1, r1=runif(8))
cs1
cs2<-transform(cs2, r2=runif(8))
cs2
cs1_sort<-cs1[order(cs1$r1),]
cs1_sort
cs2_sort<-cs2[order(cs2$r2),]
cs2_sort

```

## 付録4 survivalパッケージによるデータ分析で入力，実行したコマンド

```

library(survival)
setwd("c:/test")
getwd()
data1<-read.delim("data1.txt")
data1
clogout1<-clogit(RES~ASC+PRICE+TRANS+SOIL+strata(STR), data=data1)
clogout1
data2<-read.delim("data2.txt")
data2$TRANS_SIZE<-data2$TRANS*data2$SIZE
data2$SOIL_SIZE<-data2$SOIL*data2$SIZE
data2
clogout2<-clogit(RES~ASC+PRICE+TRANS+SOIL+TRANS_SIZE+SOIL_SIZE+strata(STR), data=data2)
clogout2

```

## 付録5 関数plot()による作図で入力，実行したコマンド

```

prb<-function(x) exp(3.56470-0.00351*x+3.17271+3.54930)/(exp(3.56470-0.00351*x+3.17271+3.54930)+exp(0))
plot(prb(1000:4000), 1000:4000, type="l", xlab="Choice Probability", ylab="PRICE")
plot(100*prb(1000:4000)*3*20, 1000:4000, type="l", xlab="Demand", ylab="PRICE")

```