

二段階二肢選択 CVM における提示額数・配布部数の選択*

寺 脇 拓

1 はじめに

近年、公共事業の全面見直しに向け、その効率性、透明性の一層の確保が求められてきている。現在では、2000年度以降の新規事業についてはその費用対効果が公表されることになっており、また2001年度からは、予算配分の見直しに向け、関連各省庁で費用対効果分析について共通の手法が導入されることが合意されている。農業分野においても、1998年12月に決定された「農政改革大綱」において、「生産基盤整備に加え、農村生活環境整備についても費用対効果分析を順次導入」することが述べられており、費用対効果分析の一層の活用が求められてきている。

これまで、農業農村整備事業のうち、土地改良法に基づく事業については、その目的が農業の生産性の向上にあり、効果の算定が比較的容易であることから、「土地改良の経済効果」と呼ばれる農林水産省構造改善局長通達¹⁾のもと、長年にわたって費用対効果分析が行われ、成果を上げてきたという経緯がある。しかし、「農政改革大綱」において費用対効果分析の対象として新たに含められた農村生活環境整備関連事業については、その目的が地域の生活環境の改善・向上とあったところにあり、費用対効果分析を実施するためには、そうした市場では評価されない外部経済効果を適切に評価することが必要となってくる。1999年度からは農業集落排水事業において、2000年度からは地域用水環境整備事業（旧：水環境整備事業）において、2001年度からは農村総合整備事業、中山間地域総合整備事業において、それぞれ費用対効果分析が試行的に導入されたが、それらの事業においては、こうした外部経済効果を評価するための手法として、CVM、特に質問形式として二段階二肢選択形式²⁾を用いたCVMが採用されている。

二段階二肢選択CVMにおけるWTP分布の推定法としては、大きく分けてパラメトリック推定法とノンパラメトリック推定法の二つがある。パラメトリック推定法は、誤差項の分布としてパラメトリックな分布型を仮定した上で、それに依存したパラメトリックなWTP分布を推定する手法である³⁾。しかし、この手法は、分布型をパラメトリックなものに規定するという点が制約的である。寺脇[14]はこの点を指摘し、二段階二肢選択CVMにおけるノンパラメトリック推定法を提案した。真の分布型は当然未知であるため、こうしたノンパラメトリック推定法は、分

*本稿を纏めるにあたって、農村環境整備センターの秋山克彦さん、木村茂基さん、農林水産省の勝部剛さんに多大なるご協力をいただきました。また、2001年度日本農業経済学会大会、及び神戸大学経済学研究科、竹内憲司先生主催の環境経済学研究会では、多くの方からの建設的なコメントを頂きました。記して感謝の意を表します。

布型をある族に規定せずに、WTP 分布を自由に推定できる点で、パラメトリック推定法よりも優れているといえる。

さらに、寺脇 [14] の提案したノンパラメトリック推定法には次の利点もある。それは電卓レベルでの計算が可能となることである。パラメトリック推定法においては、通常統計解析専用のソフトウェアが必要となり、しばしば統計学の専門家による計算が要求されるが、ノンパラメトリック推定法の計算は、統計学の専門家でない行政担当者にも可能であり、事業評価を迅速に進めることができる。また、陽表的に得られない解を最適化アルゴリズムを用いて求めるといって、専門家でない人にはある種ブラックボックスとなっている部分がないため、一般の人でも、データさえ適切に公開されていれば、計算結果の検証ができ、戦略的に評価結果を操作することができないという利点もある。こうした利点から、実際、地域用水環境整備事業では、寺脇 [14] の提案したノンパラメトリック推定法が採用されており、計算も県、あるいは市町村レベルで行われている。⁴⁾

しかし、ノンパラメトリック推定法には、提示額における生存確率しか推定することができず、またその間の仮定（Kiström [14] に従えば線形）が制約的だという批判もある。この問題点ゆえに、直感的には、ノンパラメトリック推定法で WTP 分布を推定する場合には、提示額数を多く用意することが必要だと思われる。しかし提示額数を多くすれば、それだけ調査票の印刷コストが高くなるだけでなく、標本サイズが一定のもとでは、各提示額の生存確率の推定に利用できるサブサンプルが小さくなり、評価額推定量の分散が大きくなることが予想される。もちろん、標本サイズを大きくすることによって、その問題を解消することは可能であるが、当然それは印刷コスト、配布コストの上昇を引き起こす。そこで本稿では、提示額数と標本サイズの組み合わせを変えて、モンテカルロ実験を行い、評価額推定量の平均二乗誤差（Mean Squared Error: MSE）の挙動をみることで、ノンパラメトリック推定法における提示額数・標本サイズについての望ましい調査設計を提案することに取り組む。

本稿の構成は次のようである。まず第 2 節では、二段階二肢選択 CVM におけるノンパラメトリック推定法について整理する。第 3 節では、本稿で行うモンテカルロ実験計画について述べる。第 4 節では、そのモンテカルロ実験を通して、提示額数と標本サイズがどのように評価額の誤差に影響を及ぼすかを明らかにし、ノンパラメトリック推定法における提示額数・標本サイズについての望ましい調査設計を提案する。第 5 節では、本稿の結論と今後の課題を述べる。

2 二段階二肢選択 CVM におけるノンパラメトリック推定法

二段階二肢選択形式は、Hanemann [4] や Carson [3] によって初めて提案されたものであり、その後、多くの CV 研究で適用されている。この質問形式では、まず始めに予め設定したいくつかの提示額のうちの一つを被験者に提示し、彼らがその提示額を支払う意志があるかどうかを尋ねる。そして、一般に支払う意志がある被験者に対してはさらに高い金額を、支払う意志がない被験者にはさらに低い金額を提示し、続けて支払意志を尋ねる。寺脇 [14] は、Kiström [6] の（一段階）二肢選択 CVM におけるノンパラメトリック推定法をこの二段階二肢選

提示額	B^1	B^2	B^3	...	B^k	...
yes と答えた被験者の観測値数	m^1	m^2	m^3	...	m^k	...
B^k に直面した被験者の観測値数	n^1	n^2	n^3	...	n^k	...

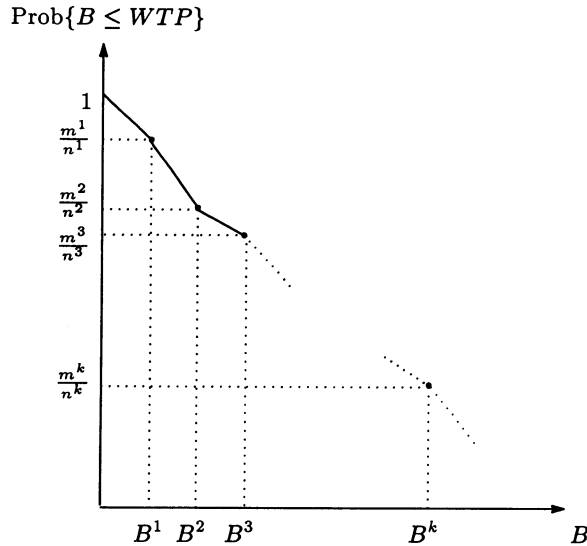


図1：ノンパラメトリック生存関数

択形式のケースに拡張し、その手法により推定される WTP の平均値、あるいは中央値と、パラメトリック推定法により推定されるそれらとの差の大きさをみることによって、WTP 分布のパラメトリックな仮定の危険性の程度を客観的に明らかにした。以下、この手法について整理する。

まず、Kriström [6] の二肢選択 CVM におけるノンパラメトリック推定法について述べよう。いま、ある提示額 B^k に直面する被験者の観測値数を n^k 、そのときに yes と答える被験者の観測値数を m^k とする。このとき、各提示額に直面する被験者のグループが同質であれば、彼らの回答はそれぞれ独立なベルヌーイ試行とみなすことができ、 m^k/n^k はその提示額 B^k に対して yes と答える確率 $\text{Prob}\{B^k \leq WTP\}$ の最尤推定量となる⁵⁾。全ての提示額に対して yes と答える確率の推定値を計算して、それらを B – $\text{Prob}\{B \leq WTP\}$ 平面にプロットし、隣り合う各点を結ぶことによって、WTP の生存関数を推定することができる（図1）。但し、全ての提示額 B^k ($k=1, \dots, n$) に対する m^k/n^k の数列が単調非増加の数列でない場合、例えばある提示額 B^k 、 B^{k+1} について $m^k/n^k < m^{k+1}/n^{k+1}$ となる場合には、それぞれの提示額に対して yes と答える確率の推定値を $(m^k + m^{k+1}) / (n^k + n^{k+1})$ で置き換え、数列が単調となるまで繰り返す。これは Ayer *et al.* [2] が提案したアルゴリズムであり、一般に PAVA (Pool Adjacent Violator Algorithm) と呼ばれる⁶⁾。平均値は、こうして推定された生存関数の下側の面積によって表され、中央値は、 $\text{Prob}\{B \leq WTP\} = 0.5$ と生存関数との交点の B 座標で表されることになる。

寺脇 [14] は、Kriström [6] における、ある提示額に直面した複数の被験者の支払回答群をそれぞれ独立なベルヌーイ試行として捉えるというアイデアを援用して、そのノンパラメトリ

ック推定法を二段階形式に拡張させた。その手法の基本的な手順は次のようである。まず、それぞれの提示額について、全ての被験者のうち、その金額の「生存（支払っても良い）」・「死亡（支払いたくない）」の判定資料となりうるもの、すなわちその提示額に直面しうるものを選定する。二段階形式においては、一人の被験者が、初期提示額だけでなく、二段階目のより高い提示額、より低い提示額の合計三つの金額に直面しうるため、一人の被験者の支払回答は、同時に三つの提示額の「生存」・「死亡」の情報をもつことになる。標本が無作為抽出されていれば、そうして選ばれた各提示額に対する被験者の回答群は、その提示額の「生存」・「死亡」についての独立なベルヌーイ試行とみなすことができる。そこで次に、選定された被験者のうち、その提示額を支払っても良いと回答した、あるいは論理的に考えて支払っても良いと答えるであろう被験者の占める割合を計算する。それが、その提示額に対する生存確率の推定値となる。

ところで、これまで、二段階二肢選択 CVM における WTP 生存関数のノンパラメトリック推定法としては、生存分析の分野で開発された Turnbull [15] の手法が代表的に援用されてきた。⁷⁾ ここで述べる寺脇 [14] のノンパラメトリック推定量は、Turnbull [15] の推定量と一般には一致しない。このことは、Turnbull [15] の推定量が全ての提示額の生存確率パラメータベクトルについての最尤推定量であるのに対し、寺脇 [14] の推定量はある提示額の生存確率パラメータについての最尤推定量であることによる。しかし、それでも両者が一致するケースもある。以下では、ある被験者のグループに提示された初期提示額と、別の被験者のグループに提示された二段階目の提示額とが一致しないケース A と、一致するケース B とに分けて、寺脇 [14] の手法をより詳しく説明する。結論を先に言えば、ケース A においては、両推定量は一致し、ケース B においては一致しない。

まずケース A について考える。いま k 番目のサブサンプルにおいて提示された初期提示額を B^k とし、yes と答えたときに提示されるより高い金額を B^{ku} 、no と答えたときに提示されるより低い金額を B^{kd} とする。そしてこのときに、順に (no, no) 、 (no, yes) 、 (yes, no) 、 (yes, yes) と答えた被験者の観測値数を m_{nn}^k 、 m_{ny}^k 、 m_{yn}^k 、 m_{yy}^k 、その総数を $n^k (= m_{nn}^k + m_{ny}^k + m_{yn}^k + m_{yy}^k)$ とする。このとき、図 2 に表されているように、 B^{kd} よりも WTP が大きい被験者の観測値数は $m_{ny}^k + m_{yn}^k + m_{yy}^k$ 、 B^k よりも WTP が大きい被験者の観測値数は、 $m_{yn}^k + m_{yy}^k$ 、 B^{ku} よりも WTP が大きい被験者の観測値数は、 m_{yy}^k である。従って、それぞれの確率の推定量は、次の (2.1) 式で表される。

$$\begin{aligned} \text{Prob} \{B^{kd} \leq WTP\} &= \frac{m_{ny}^k + m_{yn}^k + m_{yy}^k}{n^k} \\ \text{Prob} \{B^k \leq WTP\} &= \frac{m_{yn}^k + m_{yy}^k}{n^k} \\ \text{Prob} \{B^{ku} \leq WTP\} &= \frac{m_{yy}^k}{n^k} \end{aligned} \quad (2.1)$$

この推定量は、全ての生存確率パラメータベクトルについての最尤推定量である。このことを以下に示そう。まず、説明の簡単化のために、各提示額に対する生存確率を次のように表記する。

$$\begin{aligned} \text{Prob} \{B^{kd} \leq WTP\} &:= P^{kd} \\ \text{Prob} \{B^k \leq WTP\} &:= P^k \end{aligned} \quad (2.2)$$

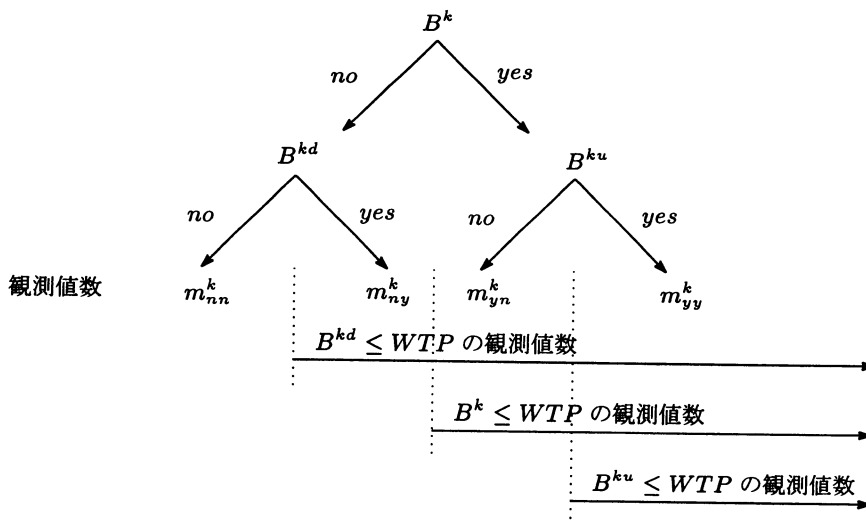


図2：各提示額に対して「支払っても良い」と答える被験者の観測値数（ケース A）

$$\text{Prob}\{B^{ku} \leq WTP\} := P^{ku}$$

このとき、対数尤度関数は次式で表される。

$$\ln L = \sum_{k=1}^M \{m_{nn}^k \ln(1-P^{kd}) + m_{ny}^k \ln(P^{kd}-P^k) + m_{yn}^k \ln(P^k-P^{ku}) + m_{yy}^k \ln P^{ku}\} \quad (2.3)$$

なお、 M はサブサンプルの数を表している。そして、尤度方程式は次式で表される。

$$\frac{\partial \ln L}{\partial P^{kd}} = -\frac{m_{nn}^k}{1-P^{kd}} + \frac{m_{ny}^k}{P^{kd}-P^k} = 0 \quad \text{for } k=1, \dots, M \quad (2.4)$$

$$\frac{\partial \ln L}{\partial P^k} = -\frac{m_{ny}^k}{P^{kd}-P^k} + \frac{m_{yn}^k}{P^k-P^{ku}} = 0 \quad \text{for } k=1, \dots, M \quad (2.5)$$

$$\frac{\partial \ln L}{\partial P^{ku}} = -\frac{m_{yn}^k}{P^k-P^{ku}} + \frac{m_{yy}^k}{P^{ku}} = 0 \quad \text{for } k=1, \dots, M \quad (2.6)$$

この連立方程式を解くことによって、(2.1) 式が得られる。

次にケース B を考える。ここでは、 k 番目のサブサンプルにおける初期提示額 B^k が、 $k-1$ 番目のサブサンプルにおける二段階目のより高い提示額 $B^{(k-1)u}$ と一致し、また $k+1$ 番目のサブサンプルにおける二段階目のより低い提示額 $B^{(k+1)d}$ と一致するようなケースを想定する。こうした提示額の設計は、二段階二肢選択形式を用いた多くの CV 研究で採用されている。

ケース B においては、ある k 番目のサブサンプルにおける初期提示額 B^k の「生存」・「死亡」の情報をもつ観測値の集合は、初期提示額 B^k に直面する被験者の観測値群、初期提示額 B^{k-1} に直面する被験者の観測値群、初期提示額 B^{k-1} に直面する被験者の観測値群をあわせたものとなる。従って、その総数は、 $n^{k-1} + n^k + n^{k+1}$ となる。そして、そのうち、 B^k よりも WTP が大きい被験者の観測値数は、図 3 に表されているように、 $m_{yy}^{k-1} + (m_{yn}^k + m_{yy}^k) + (m_{ny}^{k+1} + m_{yn}^{k+1} + m_{yy}^{k+1})$ となり、結果として、その生存確率の推定量は、次の (2.7) 式で表される。

$$\text{Prob} \{B^k \leq WTP\} = \frac{m_{yy}^{k-1} + (m_{yn}^k + m_{yy}^k) + (m_{ny}^{k+1} + m_{yn}^{k+1} + m_{yy}^{k+1})}{n^{k-1} + n^k + n^{k+1}} \quad (2.7)$$

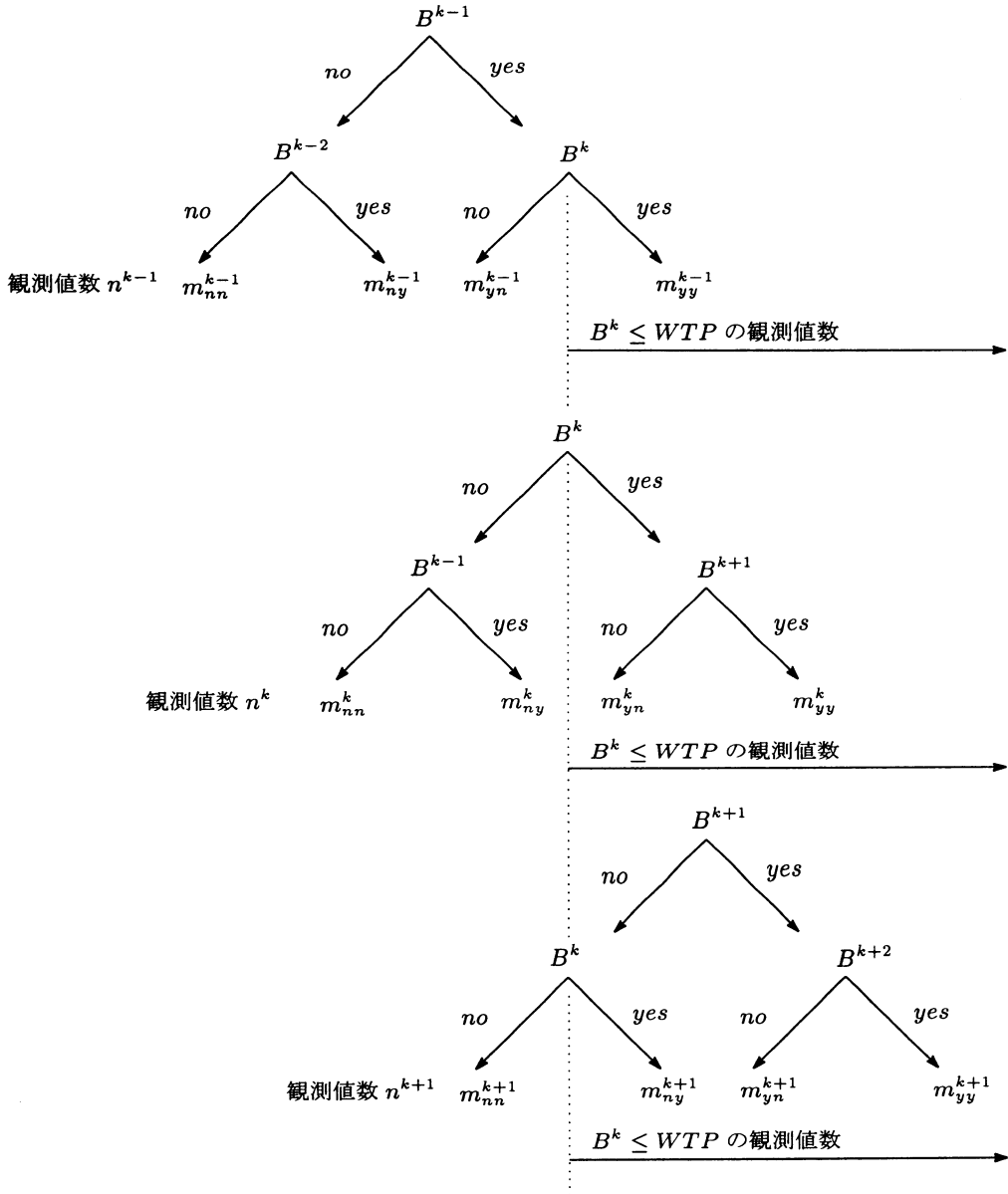


図 3：各提示額に対して「支払っても良い」と答える被験者の観測値数（ケース B）

当然，最も小さい初期提示額 B^1 ，その二段階目のより低い提示額 B^{1d} ，および最も大きい初期提示額 B^M ，その二段階目のより高い提示額 B^{Mu} の生存確率については，(2.7) 式をそのまま利用することはできない。これらの推定量は，次式で表される。

$$\text{Prob} \{B^1 \leq WTP\} = \frac{(m_{yn}^1 + m_{yy}^1) + (m_{ny}^2 + m_{yn}^2 + m_{yy}^2)}{n^1 + n^2} \quad (2.8)$$

$$\text{Prob} \{B^{1d} \leq WTP\} = \frac{m_{ny}^1 + m_{yn}^1 + m_{yy}^1}{n^1} \quad (2.9)$$

$$\text{Prob} \{B^M \leq WTP\} = \frac{m_{yy}^{M-1} + (m_{yn}^M + m_{yy}^M)}{n^{M-1} + n^M} \quad (2.10)$$

$$\text{Prob} \{B^{Mu} \leq WTP\} = \frac{m_{yy}^M}{n^M} \quad (2.11)$$

こうしたケース B においては、Turnbull [15] の手法では、推定量を陽表的に表すことができず、最適化計算が必要となるのに対し、寺脇 [14] の手法では、上記のように推定量を簡単な式で表すことができる。これが、寺脇 [14] のノンパラメトリック推定量の最大の強みである。⁸⁾

こうして推定された全ての提示額に対する生存確率を、Kriström [6] 同様、 $B - \text{Prob} \{B \leq WTP\}$ 平面にプロットし、隣り合う各点を結ぶことによって、WTP の生存関数を推定することができる。但し、全ての提示額 B^k ($k=1, \dots, n$) に対する生存確率の数列が、必ずしも単調非増加の数列となるとは限らない、例えばある提示額 B^k , B^{k+1} について、

$$\begin{aligned} & \frac{m_{yy}^{k-1} + (m_{yn}^k + m_{yy}^k) + (m_{ny}^{k+1} + m_{yn}^{k+1} + m_{yy}^{k+1})}{n^{k-1} + n^k + n^{k+1}} \\ & < \frac{m_{yy}^k + (m_{yn}^{k+1} + m_{yy}^{k+1}) + (m_{ny}^{k+2} + m_{yn}^{k+2} + m_{yy}^{k+2})}{n^k + n^{k+1} + n^{k+2}} \end{aligned} \quad (2.12)$$

となるケースが考えられる。この場合には、通常は PAVA が適用されるが、ここでのケースにおいては、単純にそれを適用するわけにはいかない。なぜなら、ある提示額 B^k , B^{k+1} の生存確率の推定量の双方に使われる観測値があり、通常の PAVA のアルゴリズムでは、標本の独立性が満たされなくなるからである。ここでは、以下の (2.13) 式に表されるように、提示額 B^k , あるいは B^{k+1} の「生存」・「死亡」の情報をもつ観測値の総数に占める、 B^k , あるいは B^{k+1} に対して支払っても良いと回答した被験者の観測値数の割合で置き換える。

$$\begin{aligned} \text{Prob} \{B^k \leq WTP\} &= \text{Prob} \{B^{k+1} \leq WTP\} \\ &= \frac{m_{yy}^{k-1} + (m_{yn}^k + m_{yy}^k) + (m_{ny}^{k+1} + m_{yn}^{k+1} + m_{yy}^{k+1}) + (m_{ny}^{k+2} + m_{yn}^{k+2} + m_{yy}^{k+2})}{n^{k-1} + n^k + n^{k+1} + n^{k+2}} \end{aligned} \quad (2.13)$$

3 モンテカルロ実験計画

本稿では、以下に述べるモンテカルロ実験を通して、上述のノンパラメトリック推定法のもとで、提示額数、標本サイズをどのように設計すればよいのかという疑問に答える。⁹⁾ まず本節では、この実験計画について整理する。

真の分布の設定 WTP の観測値は、以下に設定する真の分布から無作為に抽出されるものとする。まず、可能な限り実際の WTP の分布に近いものを真の分布として設定するという観点に立ち、寺脇 [14] において推定した農業の公益的機能に対する WTP 分布を真の分布として設定する。ここでは、WTP は対数ロジスティック分布に従うものとし、次のような密度関数をもつ

のとする。

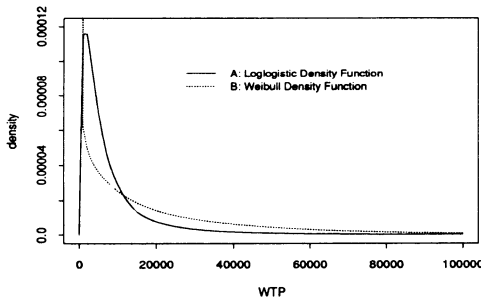
$$f(WTP) = \begin{cases} \frac{1.4065 \exp(-12.027) WTP^{0.4065}}{\{1 + \exp(-12.027) WTP^{1.4065}\}^2} & \text{if } WTP > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

以下では、この分布を A とする。この分布の平均値は14,656、中央値は5,172である。

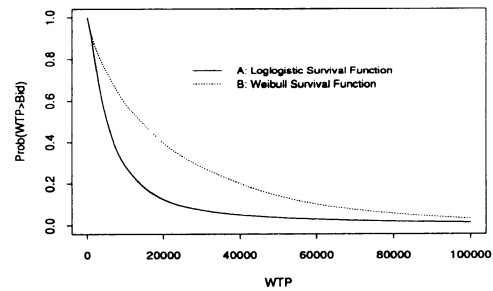
さらに、もう一つ形状の異なる分布を真の分布として設定する。WTP は単一のモードをとらないワイブル分布に従うものとし、次のような密度関数をもつものとする。

$$f(WTP) = \begin{cases} \frac{0.8}{22026.47} \left(\frac{WTP}{22026.47} \right)^{-0.2} \exp \left\{ - \left(\frac{WTP}{22026.47} \right)^{0.8} \right\} & \text{if } WTP > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

以下では、この分布を B とする。ワイブル分布の平均値は24,956、中央値は13,931である。A, B 両分布の密度関数、生存関数を図示すると図4のようになる。密度関数をみると、A 分布が単峰形を成すのに対し、B 分布はそうではないのがわかる。また生存関数をみると、A, B 両分布とも、100,000円辺りではほぼ確率が0に収束するが、その生存確率の低下のパターンには明らかな差異があることが分かる。



(a) 密度関数



(b) 生存関数

図4：設定された真の分布

提示額設計 初期提示額の数 M としては、2, 4, 6, 8 の四種類を用意する。また二段階目の提示額については、 k 番目の初期提示額を B^k 、その二段階目のより高い提示額を $B^{k\mu}$ 、より低い提示額を B^{kd} とすると、 $B^{k\mu} = B^{k+1}$ 、 $B^{kd} = B^{k-1}$ となるように設定する。但し、最大初期提示額と、最小初期提示額には、それより高い、あるいはそれより低い初期提示額が存在しないので、別途二段階目の提示額を用意する。従って、二段階目の提示額も含めると、被験者に提示される金額の数は、初期提示額の数 M のとき、 $M+2$ となる。提示額は、 $M+3$ 等確率分位点を用いて設計するものとし、最大金額、最小金額を除いた M 個の金額が初期提示額となる。

配布数設計 初期提示額当たりの配布数は全て同一とし、標本サイズ N としては、100, 300, 500, 1,000 の四種類を用意する。

計算手順 A, B それぞれの分布のもとで、WTP を N 回復元無作為抽出し、設定された提示額をもとに二段階二肢選択形式の観測値データに変換する。そしてその観測値群をもとに、各提示

額の生存確率をノンパラメトリック推定法により推定し、その間を線形補間する¹⁰⁾。そして得られた生存関数をもとに平均値、中央値を計算する。この作業を1,000回繰り返し、平均値、中央値のバイアス、標準偏差、MSE¹¹⁾を計算する。

4 モンテカルロ実験結果

4.1 平均値についての提示額・配布数設計

まず、平均値についてのモンテカルロ実験結果を表1に示す。そして、この結果を提示額数と標本サイズを水平軸にとり、MSEを垂直軸にとった鳥瞰図で表す。図5のパネル(a), (b)は、平均値についてのMSEを図示したものである。なお、各標本サイズ、提示額数のもとのMSEの間は線形で補間されている。また、この図をより詳細に分析するため、その等高線をパネル(c), (d)、提示額数とMSEとの関係をパネル(e), (f)、標本サイズとMSEとの関係をパネル(g), (h)に示す。図6は、図5の(c)~(h)の各パネルが、(a), (b)で表される鳥瞰図をどの方向から見たものとなっているのかを示したものである。本小節では、図5をもとに、平均値の提示額数・標本サイズ設計について検討する。

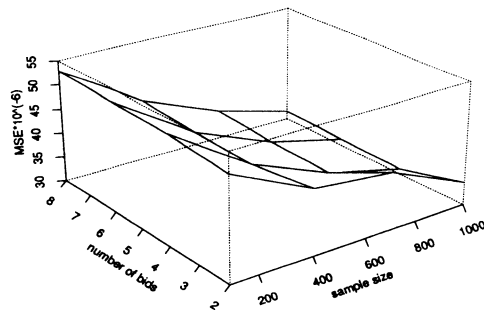
表1 平均値推定量の標準偏差・バイアス・MSE

対数ロジスティック分布												
	N=100			N=300			N=500			N=1000		
	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶
M=2	1,359	-7,002	50.871	537	-7,193	52.033	416	-7,228	52.421	288	-7,263	52.831
M=4	3,377	-5,647	43.293	1,189	-6,330	41.481	750	-6,415	41.717	447	-6,516	42.655
M=6	4,753	-4,355	41.557	2,264	-5,471	35.052	1,242	-5,811	35.311	633	-5,985	36.226
M=8	4,236	-4,098	34.739	2,899	-4,757	31.029	1,896	-5,263	31.290	1,024	-5,548	31.826
ワイブル分布												
	N=100			N=300			N=500			N=1000		
	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶
M=2	3,914	-3,710	29.083	1,611	-4,268	20.807	1,227	-4,267	19.712	851	-4,426	20.319
M=4	6,054	-1,171	38.022	2,693	-2,564	13.822	1,844	-2,905	11.840	1,206	-3,092	11.017
M=6	5,414	-1,255	30.883	3,406	-1,729	14.591	2,609	-2,066	11.074	1,571	-2,288	7.704
M=8	5,580	-728	31.666	3,779	-961	15.204	2,971	-1,251	10.393	1,772	-1,747	6.190

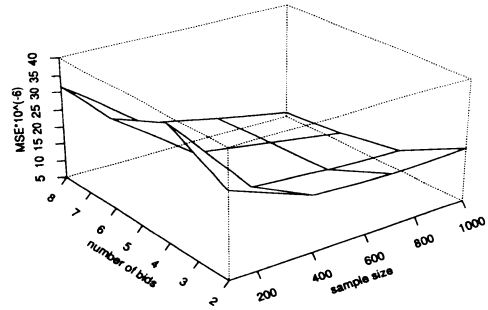
1 Mは提示額数を、Nは標本サイズを表している。

2 S.D.は標準偏差を表している。

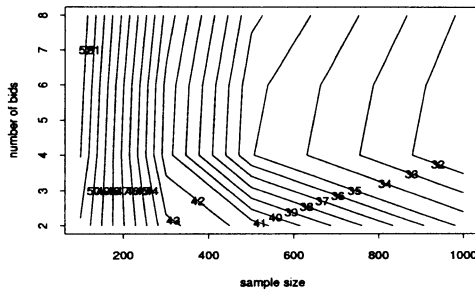
等高線を中心に考察しよう。まず、標本サイズ300までの領域では、等高線は垂直に、そして密に描かれている。これは、標本サイズ300までは、提示額の数を増やすことにはほとんど意味がなく、一方で標本サイズを大きくすることによって、急激に評価額の誤差を小さくできることを意味している。このことは、パネル(e), (f)から、標本サイズが100や300では、提示額数を増加させてもMSEがあまり変化しないこと、パネル(g), (h)から、いずれの提示額数においても、標本サイズが100から300の間では、標本サイズを大きくすることによってMSEが急激に減少する傾向を示していることから確認される。



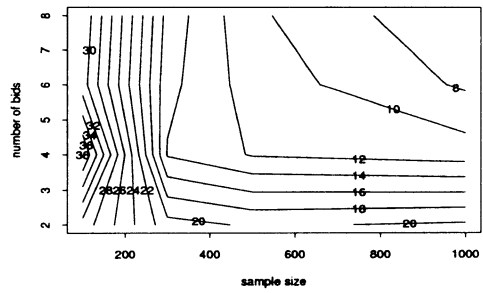
(a) 鳥瞰図 (A : 対数ロジスティック分布)



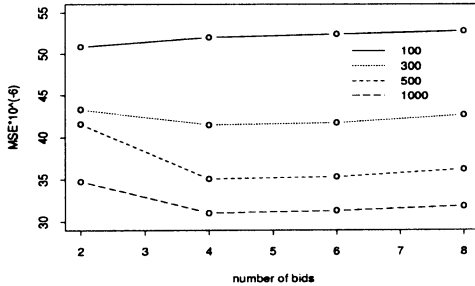
(b) 鳥瞰図 (B : ワイブル分布)



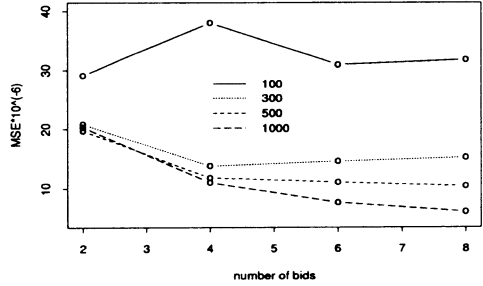
(c) 等高線 (A : 対数ロジスティック分布)



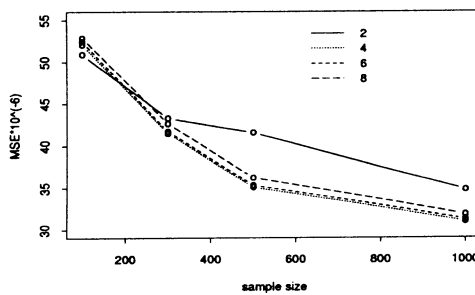
(d) 等高線 (B : ワイブル分布)



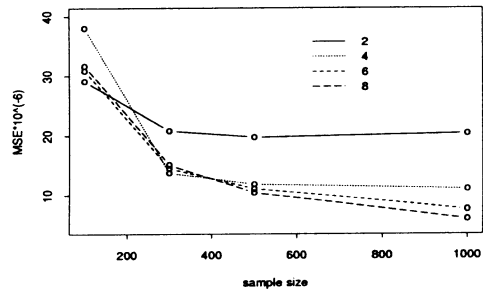
(e) 提示額数の影響 (A : 対数ロジスティック分布)



(f) 提示額数の影響 (B : ワイブル分布)



(g) 標本サイズの影響 (A : 対数ロジスティック分布)



(h) 標本サイズの影響 (B : ワイブル分布)

図 5 : 提示額数・標本サイズが及ぼす平均値推定量の MSE への影響

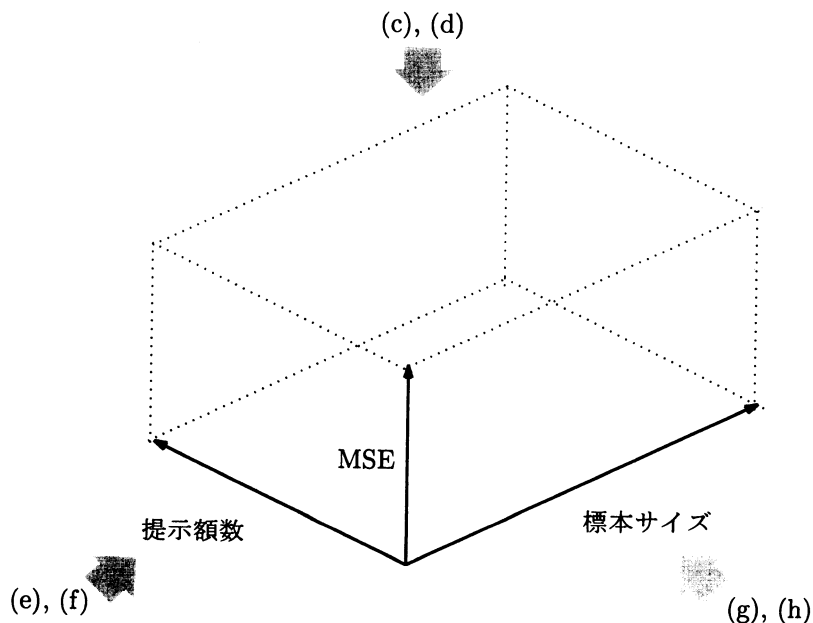


図6：MSEについての鳥瞰図（図5・図7）の見方

しかし、標本サイズが300を超えると、提示額数が4以上か以下かでこの傾向は変わってくる。提示額数が4以下の範囲では、標本サイズを大きくすることよりも、提示額を増やすことの方がより効果的であることが分かる。特にパネル(d)をみると、その範囲で等高線は水平に描かれており、ワイブル分布Bのケースでは、提示額数が2から4の範囲で標本サイズを大きくしても、MSEは全く下がらないことがわかる。またこのことは、パネル(e), (f)から、標本サイズが500や1000のケースでは、提示額数を2から4に増加させることでMSEが比較的大きく減少すること、パネル(g), (h)から、提示額数が2や4のケースでは、標本サイズを300以上に大きくしても、MSEは急激には減少しないことから確認される。

一方、提示額数が4以上の範囲では、提示額の数を増やすことにはほとんど意味がなく、標本サイズを大きくすることの方がより効果的となる傾向が保持される。しかし、等高線は100から300の範囲ほどには密に描かれておらず、標本サイズを大きくすることによる限界的な効果は減少していることがわかる。しかしそれでも、特に分布Aにおいては、パネル(g)から、提示額が4以上のケースでは、100から300までと同程度のMSEの減少傾向が、300から500の範囲で見て取れる。このことは、もし100から300に標本サイズを大きくすることに対して、標本サイズ一単位当たりの拡大コストよりもMSEの減少効果が大きいとみなされるのであれば、300に標本サイズを止めることは非効率であり、500まで大きくすることが望まれるということの意味している。

提示額数が4以上、標本サイズが500以上の領域では、分布AとBではややその傾向は異なるものの、いずれにおいてもMSEの減少の程度は緩やかになる。両分布に共通しているのは、ある提示額数において、標本サイズを大きくすることがMSEの減少につながるという当然の傾向であるが、もちろん、MSEがどれだけ減少すればよいという基準があるわけではない。ここでは、提示額数が4以上、標本サイズが500以上になれば、MSEの減少の程度は緩やかになるという結果を受けて、提示額数4、標本サイズ（すなわち有効回答部数）500を、平均値評価額を

推定する際に最低限必要とされる提示額数，標本サイズとして提案する。

4. 2 中央値についての提示額・配布数設計

表 2 は，中央値についてのモンテカルロ実験結果を示したものである。そして，この結果を鳥瞰図で図示したものが，図 7 のパネル(a), (b)である。平均値同様，パネル(c), (d)はその等高線を，パネル(e), (f)は提示額数と MSE との関係を，パネル(g), (h)は標本サイズと MSE との関係を示している。本小節では，図 7 をもとに，中央値の提示額数・標本サイズ設計について検討する。

表 2 中央値推定量の標準偏差・バイアス・MSE

対数ロジスティック分布												
	N=100			N=300			N=500			N=1000		
	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶
M=2	759	248	0.638	395	219	0.204	314	224	0.149	208	204	0.085
M=4	871	153	0.782	445	98	0.207	350	125	0.138	235	105	0.066
M=6	951	100	0.915	559	75	0.318	431	68	0.191	305	71	0.098
M=8	1,118	222	1.299	662	60	0.442	500	37	0.251	356	42	0.129

ワイブル分布												
	N=100			N=300			N=500			N=1000		
	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶	S.D.	BIAS	MSE×10 ⁻⁶
M=2	2,498	654	6.666	1,394	768	2.533	1,006	690	1,488	745	669	1.003
M=4	2,783	492	7.986	1,598	366	2.688	1,197	472	1,655	782	405	0.776
M=6	3,283	380	10.924	1,973	259	3.959	1,543	236	2,435	1,046	233	1.148
M=8	3,724	298	13.956	2,302	259	5.367	1,710	90	2,934	1,252	163	1.593

1 M は提示額数を，N は標本サイズを表している。

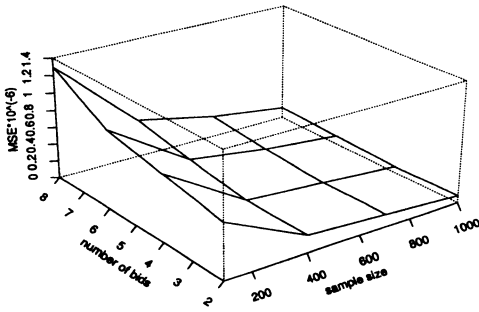
2 S.D. は標準偏差を表している。

等高線を中心に考察する。まず，平均値と決定的に違うところは，全体的に等高線が提示額数を減少させるほど低くなるように描かれているところである。このことは，パネル(e), (f)から明らかな傾向としてみる事ができる。中央値は，生存関数全体を推定する必要がなく，要は，生存確率が0.5となるところの前後の生存確率をより正確に推定することの方が求められるので，このような傾向が観察されたものと思われる¹²⁾。従って，提示額数は2で十分だと結論付けられる。

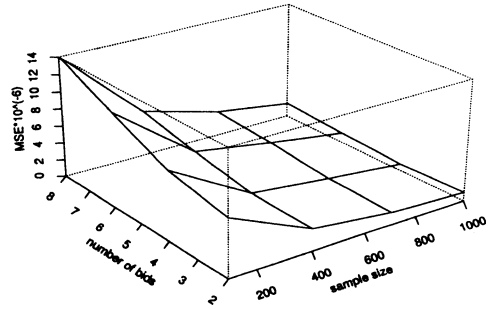
一方，標本サイズについても，全体的に等高線は標本サイズを大きくするほど低くなるように描かれている。このことは，パネル(g), (h)からも明らかである。しかし，その低下傾向は，等高線が300までの範囲では密に描かれているのに対し，それ以降は粗に描かれている。パネル(g), (h)をみても，300までは極めて急激に MSE が低下し，それ以降は緩やかな低下傾向を示している。このことから，ここでは，標本サイズ300を，中央値を推定する際に最低限必要な標本サイズとして提案する。整理すると，中央値の推定に必要な提示額数は2，標本サイズは300である。

5 むすび

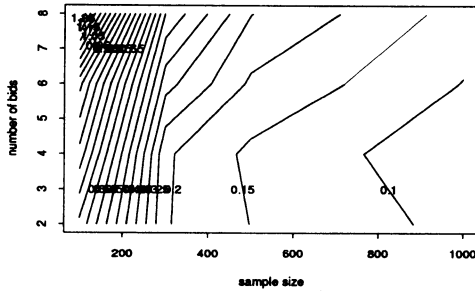
本稿では，モンテカルロ実験を通して，ノンパラメトリック推定法における提示額数・標本サイズについての望ましい調査設計を提案することに取り組んだ。真の生存関数が滑らかに逓減す



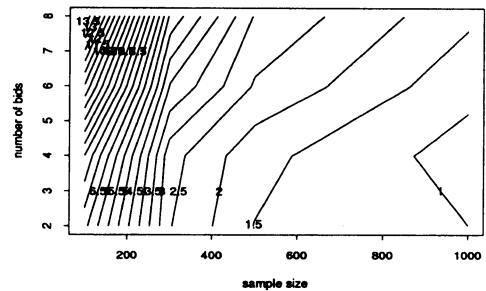
(a) 鳥瞰図 (A : 対数ロジスティック分布)



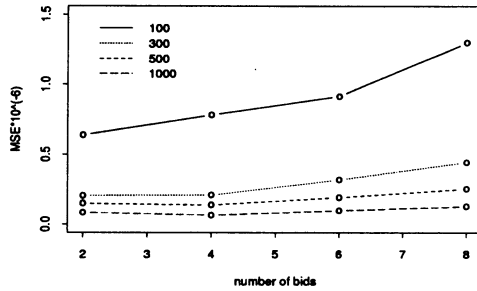
(b) 鳥瞰図 (B : ワイブル分布)



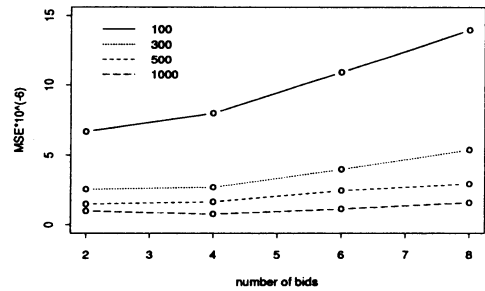
(c) 等高線 (A : 対数ロジスティック分布)



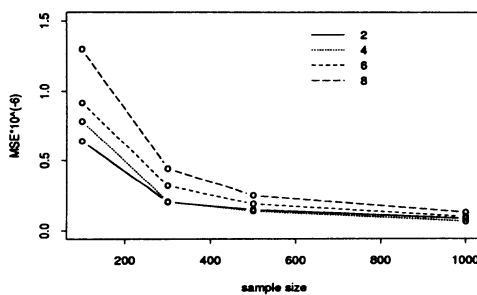
(d) 等高線 (B : ワイブル分布)



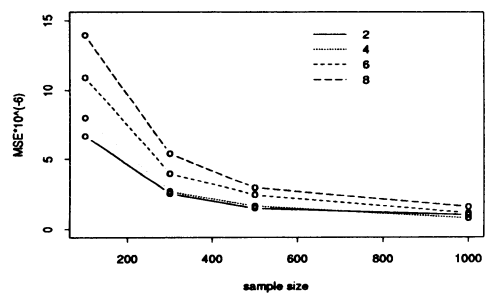
(e) 提示額数の影響 (A : 対数ロジスティック分布)



(f) 提示額数の影響 (B : ワイブル分布)



(g) 標本サイズの影響 (A : 対数ロジスティック分布)



(h) 標本サイズの影響 (B : ワイブル分布)

図7：提示額数・標本サイズが及ぼす中央値推定量のMSEへの影響

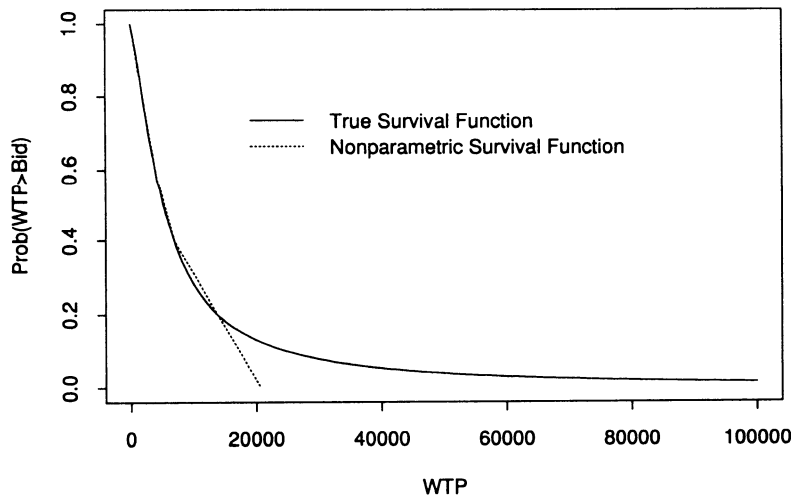


図 8：平均値を過小評価するノンパラメトリック生存関数の一例

る関数で、予め分布の形状が予想されており、その分布に基づく等確率分位点を用いた提示額設計を行うという条件のもとで、本実験から提案される調査設計は次のように要約される。

- 平均値評価額を推定する際には、初期提示額数は 4 でよく、標本サイズは最低 500 必要である。
- 中央値評価額を推定する際には、初期提示額数は 2 でよく、標本サイズは最低 300 必要である。

通常、有効回答率は 30% 前後であるため、この結果から、調査票配布部数は、平均値では 1,600 部程度、中央値では 1,000 部程度必要だということになる。¹³⁾

但し、この結論の実際の調査への適用には注意が必要である。本分析でみたのは、あくまで MSE がどのあたりで収束していくかということであり、その大きさがどれほど小さくなるかということではない。例えば、表 1 のバイアスの欄をみると、それらは軒並みマイナスの値をとっており、その大きさは、最も小さいものでも 4,000 円以上になっている。真の平均値が 14,656 円であることから、それはおよそ 30% を占める。望ましい設計として提案した、初期提示額数 4、標本サイズ 500 においては、そのバイアスは 6,415 円にものぼる。確かに MSE が収束するという意味では、この調査設計は望ましいといえるのであるが、これほどのバイアスを許容できるかといわれれば、なかなかそうはいえない。繰り返すが、本稿で提案している調査設計は、あくまで MSE の収束性でもって判断されているものであり、大きさでは判断されていないことに注意されたい。

こうした平均値推定量におけるバイアスの大きさは、モンテカルロ実験計画、特に提示額設計に改善の余地があることを暗示している。本実験では、等確率分位点による提示額設計を採用したが、この設計では、提示額数が少ないとどうしても最高提示額が小さく設定され、線形補間の仮定のもとでは推定値が過小評価されてしまう可能性が高い。図 8 は、今回の $M=2$ 、 $N=100$ のケースの実験（真の分布は対数ロジスティック分布）において 1,000 回推定された各提示額の生存確率の平均値を図示したものである。最大提示額における生存確率とその一つ前の提示額にお

る生存確率との間を結ぶ直線を外挿することで、生存確率を0にもっていく計算方法では、この図に表されるように、滑らかな生存確率の遞減傾向を推定することは難しく、結果大きなバイアスを導いてしまう。この問題を解決するためには、等確率分位点の最高提示額よりもさらに高い提示額を用意するような設計法を提案することができる。しかしその有効性については明らかではなく、今後取り組むべき課題だといえる。

また、提示額ごとの配布部数についての検討もなされていない。全ての提示額について同数配布という設計方法が果たして効率的なものかどうか、検討が必要である。これらの点を考慮した実験については、今後の課題としたい。

補論A 二段階二肢選択 CVM におけるパラメトリック推定法

本補論では、二段階二肢選択形式における WTP 分布のパラメトリック推定法について整理する。二肢選択型の CVM のモデルとしては、主に間接効用関数アプローチと評価関数アプローチがあるが、McConnel [8] によって示されているように、両モデルは同じ効用理論から導出されるものであり、決定的な関係において両者に差はない。本稿では、WTP を確率変数として捉え、その分布の仮定をもとにモンテカルロ実験を行っていることから、WTP を明示的に確率変数として定式化する評価関数アプローチについて、以下に紹介する¹⁴⁾。

まず、評価関数を次のように定式化する。

$$\log WTP_i = \mu + \varepsilon_i \quad (\text{A.1})$$

但し、 WTP_i は被験者 i の真の WTP、 μ は未知のパラメータ、 ε_i は平均 0、標準偏差 σ をもつ i. i. d. の誤差項とする。なお、WTP は正であるとして自然対数がとられている。

ここで、被験者 i に提示された始めの金額を B_i とし、この金額に対して *yes* と答えたときに提示されるより高い金額を B_i^u 、*no* と答えたときに提示されるより低い金額を B_i^d としよう。このとき、被験者 i が B_i に対して *yes* と答え、 B_i^u に対しても *yes* と答える尤度 $\pi^{uu}(B_i, B_i^u)$ 、 B_i に対して *yes* と答え、 B_i^u に対して *no* と答える尤度 $\pi^{un}(B_i, B_i^u)$ 、 B_i に対して *no* と答え、 B_i^d に対して *yes* と答える尤度 $\pi^{du}(B_i, B_i^d)$ 、 B_i に対して *no* と答え、 B_i^d に対しても *no* と答える尤度 $\pi^{dn}(B_i, B_i^d)$ は、それぞれ次のように表される。

$$\begin{aligned} \pi^{uu}(B_i, B_i^u) &= \text{Prob} \{B_i^u \leq WTP_i\} = 1 - F\{(\log B_i^u - \mu) / \sigma\} \\ \pi^{un}(B_i, B_i^u) &= \text{Prob} \{B_i \leq WTP_i < B_i^u\} = F\{(\log B_i^u - \mu) / \sigma\} - F\{(\log B_i - \mu) / \sigma\} \\ \pi^{du}(B_i, B_i^d) &= \text{Prob} \{B_i^d \leq WTP_i < B_i\} = F\{(\log B_i - \mu) / \sigma\} - F\{(\log B_i^d - \mu) / \sigma\} \\ \pi^{dn}(B_i, B_i^d) &= \text{Prob} \{WTP_i < B_i^d\} = F\{(\log B_i^d - \mu) / \sigma\} \end{aligned} \quad (\text{A.2})$$

但し、 $F(\cdot)$ は任意の分布関数を表している。 $F(\cdot)$ をそれぞれ正規分布関数、ロジスティック分布関数、ガンベル（の最小値）分布関数と仮定すれば、この定式化により、WTP は対数正規分布、対数ロジスティック分布、ワイブル分布をすることが仮定されることになる。

(A.2) 式から各被験者の尤度が得られるので、それらをもとにして対数尤度関数をつくり、最尤法によってパラメータベクトル $\theta := (\mu, \sigma)$ を推定する。対数尤度関数は次のように表される。

$$\begin{aligned} \ln L(\theta) = \sum_{i=1}^N \{ & d_i^{yy} \ln \pi^{yy}(B_i, B_i^y) + d_i^{yn} \ln \pi^{yn}(B_i, B_i^y) \\ & + d_i^{ny} \ln \pi^{ny}(B_i, B_i^d) + d_i^{nn} \ln \pi^{nn}(B_i, B_i^d) \} \end{aligned} \quad (\text{A.3})$$

$d_i^{yy}, d_i^{yn}, d_i^{ny}, d_i^{nn}$ は、それぞれ (yes, yes), (yes, no), (no, yes), (no, no) と答えたときに 1, それ以外のときに 0 をとる二値定義変数, N は観測値数を表している。最尤推定量 $\hat{\theta}$ は、尤度方程式 $\partial \ln L(\theta) / \partial \theta = 0$ の解である。

パラメータ推定値が得られれば、それらを $1 - F\{(\log WTP - \mu) / \sigma\}$ に代入することで、WTP の生存関数が得られ、そこから平均値、中央値を計算することができる。まず平均値については、それは一般に生存関数 $S(WTP)$ の下側の面積で計られ、次式で計算される。

$$E(WTP) = \int_0^\infty S(WTP) dWTP = \int_0^\infty 1 - F\{(\log WTP - \mu) / \sigma\} dWTP \quad (\text{A.4})$$

但し、生存関数が 0 に収束せず、平均値が発散してしまう場合がある。この点に配慮し、また控えめな評価額を導出するという観点から、しばしば、ある最大金額 WTP_{\max} で分布を切断したときの切断平均値 (Truncated Mean) が採用される。これは次式で表される。

$$E(WTP) = \int_0^{WTP_{\max}} \frac{S(WTP)}{1 - S(WTP_{\max})} dWTP \quad (\text{A.5})$$

一方中央値は、 $S(WTP) = 0.5$ を解くことによって計算されるが、(A.2) 式において仮定される分布が 0 について対称なものであれば、 $\exp(\mu)$ で計算される。

補論 B S-PLUS 関数

本補論では、ここでのモンテカルロ実験の際に作成した S-PLUS¹⁵⁾ の関数を紹介し、その使い方を述べる。¹⁶⁾ ここで提供する関数により、提示額の数を変えたとき、そして標本サイズを変えたときに、評価額推定量のバイアス、分散、MSE がどのように変化するかをみることができる。また、当関数はノンパラメトリック推定法だけでなく、補論 A で述べたパラメトリック推定法による評価額推定も併せて行うようプログラミングされており、両推定法のパフォーマンスを比較することもできる。但し、次の制約がある。

- 真の分布は、ワイブル分布か対数ロジステック分布のいずれかに限られる。
- 二段階目の提示額については、 k 番目の初期提示額を B^k 、その二段階目のより高い提示額を B^{ku} 、より低い提示額を B^{kd} とすると、 $B^{ku} = B^{k+1}$ 、 $B^{kd} = B^{k-1}$ となるように設定される。
- 初期提示額数を M として、 $M+2$ 個の提示額は $M+3$ 等確率分位点を用いて設計される。¹⁷⁾
- 初期提示額当たりの配布数は全て同一と設定される。例えば、標本サイズを 100、初期提示額数を 4 とするとき、初期提示額当たりの配布数は、全ての初期提示額について 25 となる。

作業手順は次のようである。

1. 表 3 に表される関数群を定義する。
2. dbdc. mont. exp. 1 (M, N, B, model, p. gamma, p. lambda, trun. v) の関数によりモンテカルロ実験を実行し、結果として作成されるデータセットを S オブジェクトに格納する。各引数の意味は次の通り、

- M : 反復数

- N : 標本サイズ
- B : 初期提示額の数（最大18）。
- model : 真の分布をワイブル分布と仮定するとき 1 , 対数ロジスティック分布と仮定するとき 2 。
- p. gamma, p. lambda : 仮定される分布のパラメータ

—ワイブル分布の密度関数

$$f(WTP) = \frac{p. gamma}{p. lambda} \left(\frac{WTP}{p. lambda} \right)^{p. gamma - 1} \exp \left\{ - \left(\frac{WTP}{p. lambda} \right)^{p. gamma} \right\}$$

—対数ロジスティック分布の密度関数

$$f(WTP) = \frac{p. gamma \exp(-p. lambda) WTP^{p. gamma - 1}}{\{1 + \exp(-p. lambda) WTP^{p. gamma}\}^2}$$

- trun. v : ノンパラメトリック分布の切断値。当然，最高提示額よりも大きな値として設定される必要がある。切断しない場合は Inf とすればよいが，ノンパラメトリック生存関数が 0 に収束しないケースも十分考えられ，その場合平均値が無限大となってしまうことに注意されたい。また，併せて推定されるパラメトリック生存関数は，ノンパラメトリック生存関数が 0 となるところの金額で切断されるよう設定されており，必ずしもここで設定される切断値で分布が切断されるわけではないことにも注意されたい。

例えば，第 3 節で述べた，形状パラメータ 0.8，尺度パラメータ 22026.47 のワイブル分布を仮定し，初期提示額数を 4，標本サイズ 100，切断値を 100000 としたときの実験を 10 回繰り返すとき，そしてその結果を S データフレームオブジェクト wei. 4.100 に格納するとき，コマンドは次のようになる。

```
wei. 4. 100 <- dbdc. mont. exp. 1 (10, 100, 4, 1, 0.8, 22026.47, 100000)
```

3. 最後に，result. dbdc. mont. exp. 1 (X, model, p. gamma, p. lambda) の関数により，モンテカルロ実験結果を要約する。各引数の持つ意味は次の通り，

- X : dbdc. mont. exp. 1 により作成された S データフレームオブジェクト
- model : 真の分布をワイブル分布と仮定するとき 1，対数ロジスティック分布と仮定するとき 2 。
- p. gamma, p. lambda : 仮定される分布のパラメータ。dbdc. mont. exp. 1 と同様。

結果，ノンパラメトリックモデル (Nonp)，対数ロジスティックモデル (logis)，対数正規モデル (norm)，ワイブルモデル (Wei) ごとに，平均値推定量の平均値 (mean. mean)，分散 (mean. var)，標準偏差 (mean. s. d.)，バイアス (mean. bias)， 10^6 で除した MSE (mean. mse-6)，そして中央値推定量の平均値 (median. mean)，分散 (median. var)，標準偏差 (median. s. d.)，バイアス (median. bias)， 10^6 で除した MSE (median. mse-6) がそれぞれ計算される。

表 3 S-PLUS 関数

```
1 : ##### S-PLUS functions for analysing the influence of sample size
2 : ##### and the number of bids on the error of WTP estimation
3 : ##### in double-bounded dichotomous choice CVM
4 : ##### (C) 2001 Taku TERAWAKI
5 : ##### E-MAIL: masala.terawaki@nifty.ne.jp
6 : ##### URL: http://www.taku-t.com/
```

```

7 : ##### likelihood functions in parametric model #####
8 : ### log-logistic dist. ###
9 : DBDCL.obj <- function(r, X)
10 : {
11 :     W2H <- (log(X$BIDU) - r[1])/r[2]
12 :     W <- (log(X$BID) - r[1])/r[2]
13 :     W2L <- (log(X$BIDD) - r[1])/r[2]
14 :     P.yy <- log(1 - plogis(W2H))
15 :     P.yn <- log(plogis(W2H) - plogis(W))
16 :     P.ny <- log(plogis(W) - plogis(W2L))
17 :     P.nn <- log(plogis(W2L))
18 :     obj <- X$Y1 * X$Y2 * P.yy + X$Y1 * (1 - X$Y2) * P.yn + (1 - X$Y1) *
19 :         X$Y2 * P.ny + (1 - X$Y1) * (1 - X$Y2) * P.nn
20 :     - sum(obj)
21 : }
22 : ### log-normal dist. ###
23 : DBDCP.obj <- function(r, X)
24 : {
25 :     W2H <- (log(X$BIDU) - r[1])/r[2]
26 :     W <- (log(X$BID) - r[1])/r[2]
27 :     W2L <- (log(X$BIDD) - r[1])/r[2]
28 :     P.yy <- log(1 - pnorm(W2H))
29 :     P.yn <- log(pnorm(W2H) - pnorm(W))
30 :     P.ny <- log(pnorm(W) - pnorm(W2L))
31 :     P.nn <- log(pnorm(W2L))
32 :     obj <- X$Y1 * X$Y2 * P.yy + X$Y1 * (1 - X$Y2) * P.yn + (1 - X$Y1) *
33 :         X$Y2 * P.ny + (1 - X$Y1) * (1 - X$Y2) * P.nn
34 :     - sum(obj)
35 : }
36 : ### weibull dist. ###
37 : DBDCW.obj <- function(r, X)
38 : {
39 :     W2H <- (log(X$BIDU) - r[1])/r[2]
40 :     W <- (log(X$BID) - r[1])/r[2]
41 :     W2L <- (log(X$BIDD) - r[1])/r[2]
42 :     P.yy <- log(exp(- exp(W2H)))
43 :     P.yn <- log(- exp(- exp(W2H)) + (exp(- exp(W))))
44 :     P.ny <- log(- exp(- exp(W)) + (exp(- exp(W2L))))
45 :     P.nn <- log(1 - exp(- exp(W2L)))
46 :     obj <- X$Y1 * X$Y2 * P.yy + X$Y1 * (1 - X$Y2) * P.yn + (1 - X$Y1) *
47 :         X$Y2 * P.ny + (1 - X$Y1) * (1 - X$Y2) * P.nn
48 :     - sum(obj)
49 : }
50 : ##### the functions for calculating representative values in nonparametric model #####
51 : ### mean ###
52 : nonpara.mean.wtp <- function(bids, est.prob)
53 : {
54 :     m <- length(est.prob) - 1
55 :     area <- 0
56 :     for(i in 1:m) {
57 :         area <- area + ((bids[i + 1, ] - bids[i, ]) * (est.prob[
58 :             i + 1] + est.prob[i]))/2
59 :     }
60 :     area <- as.data.frame(area)
61 :     area
62 : }
63 : ### median ###
64 : nonpara.median.wtp <- function(bids, est.prob)
65 : {
66 :     m <- length(est.prob) - 1
67 :     for(i in 1:m) {
68 :         if(est.prob[i] <= 0.5)
69 :             break

```

```

70 :             median.wtp <- ifelse(est.prob[i + 1] == est.prob[i], 1, ((
71 :                 bids[i + 1, ] - bids[i, ])/(est.prob[i + 1] -
72 :                 est.prob[i]) * (0.5 - est.prob[i])) + bids[i, ])
73 :         }
74 :         median.wtp <- as.data.frame(median.wtp)
75 :         median.wtp
76 :     }
77 : ##### survival functions in parametric model #####
78 : ### log-logistic dist. ###
79 : param.surv.l <- function(Z, Xmax, cons, sigm)
80 : {
81 :     plogis((cons - log(Z))/sigm)/(1 - plogis((cons - log(Xmax))/sigm)
82 : )
83 : }
84 : ### log-normal dist. ###
85 : param.surv.n <- function(Z, Xmax, cons, sigm)
86 : {
87 :     pnorm((cons - log(Z))/sigm)/(1 - pnorm((cons - log(Xmax))/sigm))
88 : }
89 : ### weibull dist. ###
90 : param.surv.w <- function(Z, Xmax, cons, sigm)
91 : {
92 :     exp(- exp((log(Z) - cons)/sigm))/(1 - exp(- exp((log(Xmax) -
93 :     cons)/sigm)))
94 : }
95 : ##### the function for a monte carlo experiment #####
96 : dbdc.mont.exp.1 <- function(M, N, B, model, p.gamma, p.lambda, trun.v)
97 : {
98 :     ## M : the number of iteration
99 :     ## N : sample size
100 :    ## B : the number of initial bids (which is allowed up to 18)
101 :    ## model : model=1 means weibull model, model=2 means loglogistic model.
102 :    ## p.gamma, p.lambda : the parameters in the following density functions
103 :    ## ---weibull model
104 :    ## f(x)={p.gamma/p.lambda*(x/p.lambda)^(p.gamma-1)}*{exp-(x/p.lambda)^p.gamma}
105 :    ## ---loglogistic model
106 :    ## f(x)={p.gamma*exp(-p.lambda)*x^(p.gamma-1)}/{1+exp(-p.lambda)*x^p.gamma}^2
107 :    ## trun.v : the truncation value of survival functions ; the value must be more
108 :    ## than the highest bid. also if the model has no truncation value,
109 :    ## then trun.v=Inf
110 :    cv.datum <- matrix(0, N, 6)
111 :    cv.datum <- as.data.frame(cv.datum)
112 :    var.name <- c("BIDD", "BID", "BIDU", "Y1", "Y2", "TWTP")
113 :    dimnames(cv.datum)[[2]] <- var.name
114 :    bid.vector <- matrix(0, B + 2, 1)
115 :    bid.vector <- as.data.frame(bid.vector)
116 :    mont.mean <- matrix(0, M, 4)
117 :    mont.mean <- as.data.frame(mont.mean)
118 :    mont.median <- matrix(0, M, 4)
119 :    mont.median <- as.data.frame(mont.median)
120 :    num.obs <- matrix(0, 4, B + 4)
121 :    num.obs <- as.data.frame(num.obs)
122 :    est.prob <- matrix(0, B + 4, 3)
123 :    est.prob <- as.data.frame(est.prob)
124 :    mont.est.prob <- matrix(0, M, B + 4)
125 :    mont.est.prob <- as.data.frame(mont.est.prob)
126 :    est.coef.objec <- matrix(0, M, 9)
127 :    est.coef.objec <- as.data.frame(est.coef.objec)
128 :    check.vec <- matrix(0, B + 1, 1)
129 :    check.vec <- as.data.frame(check.vec)
130 :    name.col <- c("Mean.Nonp", "Mean.logis", "Mean.norm", "Mean.Wei",
131 :    Medi.Nonp, "Medi.logis", "Medi.norm", "Medi.Wei",
132 :    a.loigs, "b.logis", "a.norm", "b.norm", "a.wei",

```

```

133 :         b.wei, "like.logis", "like.norm", "like.wei")
134 :     bids.nonp.name <- c("p1", "p2", "p3", "p4", "p5", "p6", "p7",
135 :         p8, "p9", "p10", "p11", "p12", "p13", "p14", "p15",
136 :         p16, "p17", "p18", "p19", "p20", "p21", "p22")
137 :     ## Mean.Nonp : the estimated mean by nonparametric method
138 :     ## Mean.logis : the estimated mean by parametric method with loglogistic model
139 :     ## Mean.norm : the estimated mean by parametric method with lognormal model
140 :     ## Mean.Wei : the estimated mean by parametric method with weibull model
141 :     ## Medi.Nonp : the estimated median by nonparametric method
142 :     ## Medi.logis : the estimated median by parametric method with loglogistic model
143 :     ## Medi.norm : the estimated median by parametric method with lognormal model
144 :     ## Medi.Wei : the estimated median by parametric method with weibull model
145 :     ## a.loigs, b.logis : the parameter estimates in loglogistic model
146 :     ## a.norm, b.norm : the parameter estimates in lognormal model
147 :     ## a.wei, b.wei : the parameter estimates in weibull model
148 :     ## like.logis : maximum log-likelihood in loglogistic model
149 :     ## like.norm : maximum log-likelihood in lognormal model
150 :     ## like.wei : maximum log-likelihood in weibull model
151 :     ## p1-p22 : the estimates of survival probability in nonparametric model
152 :     bids.nonp.name <- bids.nonp.name[1:(B + 4)]
153 :     name.col <- c(name.col, bids.nonp.name)
154 :     mean.logis <- p.lambda/p.gamma
155 :     var.logis <- 1/p.gamma
156 :     for(i in 1:(B + 2)) {
157 :         bid.vector[i, ] <- if(model == 1) qweibull(i/(B + 3),
158 :             p.gamma, p.lambda) else exp(qlogis(i/(B +
159 :             3), mean.logis, var.logis))
160 :     }
161 :     for(i in 1:M) {
162 :         for(j in 1:N) {
163 :             gen.elem <- rlogis(1, mean.logis, var.logis)
164 :             gen.data <- exp(gen.elem)
165 :             num.bid.vec <- ifelse(B == 1, 2, sample(2:(B + 1),
166 :                 1))
167 :             bid.d <- bid.vector[num.bid.vec - 1, ]
168 :             bid.m <- bid.vector[num.bid.vec, ]
169 :             bid.u <- bid.vector[num.bid.vec + 1, ]
170 :             cv.datum[j, 1] <- bid.d
171 :             cv.datum[j, 2] <- bid.m
172 :             cv.datum[j, 3] <- bid.u
173 :             cv.datum[j, 4] <- if(bid.m <= gen.data) 1 else 0
174 :             cv.datum[j, 5] <- if((bid.m <= gen.data & bid.u <=
175 :                 gen.data) | (bid.m > gen.data & bid.d <=
176 :                 gen.data)) 1 else 0
177 :             cv.datum[j, 6] <- gen.data
178 :         }
179 :     ## Terawaki's nonparametric model ##
180 :     for(k in 1:B) {
181 :         num.obs[1, k + 2] <- nrow(cv.datum[cv.datum[,
182 :             Y1] == 0 & cv.datum[, "Y2"] == 0 &
183 :             cv.datum[, "BID"] == bid.vector[k + 1, ],
184 :             ])
185 :         num.obs[2, k + 2] <- nrow(cv.datum[cv.datum[,
186 :             Y1] == 0 & cv.datum[, "Y2"] == 1 &
187 :             cv.datum[, "BID"] == bid.vector[k + 1, ],
188 :             ])
189 :         num.obs[3, k + 2] <- nrow(cv.datum[cv.datum[,
190 :             Y1] == 1 & cv.datum[, "Y2"] == 0 &
191 :             cv.datum[, "BID"] == bid.vector[k + 1, ],
192 :             ])
193 :         num.obs[4, k + 2] <- nrow(cv.datum[cv.datum[,
194 :             Y1] == 1 & cv.datum[, "Y2"] == 1 &
195 :             cv.datum[, "BID"] == bid.vector[k + 1, ],

```

```

196 :                               ])
197 :                               }
198 :                               est.prob[1, 1] <- 1
199 :                               est.prob[1, 2] <- 1
200 :                               est.prob[1, 3] <- est.prob[1, 1]/est.prob[1, 2]
201 :                               for(k in 1:(B + 2)) {
202 :                                   est.prob[k + 1, 1] <- num.obs[4, k] + num.obs[3,
203 :                                       k + 1] + num.obs[4, k + 1] + num.obs[2, k +
204 :                                       2] + num.obs[3, k + 2] + num.obs[4, k + 2
205 :                                       ]
206 :                                   est.prob[k + 1, 2] <- sum(num.obs[, k]) + sum(
207 :                                       num.obs[, k + 1]) + sum(num.obs[, k + 2])
208 :                                   est.prob[k + 1, 3] <- est.prob[k + 1, 1]/est.prob[
209 :                                       k + 1, 2]
210 :                               }
211 :                               est.prob[B + 4, 1] <- 0
212 :                               est.prob[B + 4, 2] <- 1
213 :                               est.prob[B + 4, 3] <- est.prob[B + 4, 1]/est.prob[B + 4,
214 :                                       2]
215 :                               for(k in 1:50) {
216 :                                   for(l in 1:(B + 1)) {
217 :                                       check.vec[l, ] <- if(est.prob[l + 1, 3] >=
218 :                                           est.prob[l + 2, 3]) 0 else 1
219 :                                   }
220 :                                   if(sum(check.vec[, 1]) == 0)
221 :                                       break
222 :                                   for(t in 2:(B + 3)) {
223 :                                       est.prob[t, 1] <- if(est.prob[t, 3] <
224 :                                           est.prob[t + 1, 3]) est.prob[t, 1] +
225 :                                           est.prob[t + 1, 1] else est.prob[t, 1
226 :                                           ]
227 :                                       est.prob[t, 2] <- if(est.prob[t, 3] <
228 :                                           est.prob[t + 1, 3]) est.prob[t, 2] +
229 :                                           est.prob[t + 1, 2] else est.prob[t, 2
230 :                                           ]
231 :                                       est.prob[t, 3] <- est.prob[t, 1]/est.prob[
232 :                                           t, 2]
233 :                                       est.prob[t + 1, 1] <- if(est.prob[t, 3] <
234 :                                           est.prob[t + 1, 3]) est.prob[t, 1] +
235 :                                           est.prob[t + 1, 1] else est.prob[t +
236 :                                           1, 1]
237 :                                       est.prob[t + 1, 2] <- if(est.prob[t, 3] <
238 :                                           est.prob[t + 1, 3]) est.prob[t, 2] +
239 :                                           est.prob[t + 1, 2] else est.prob[t +
240 :                                           1, 2]
241 :                                       est.prob[t + 1, 3] <- est.prob[t + 1, 1]/
242 :                                           est.prob[t + 1, 2]
243 :                                   }
244 :                               }
245 :                               last.bid <- if(est.prob[B + 2, 3] != est.prob[B + 3, 3]) (
246 :                                   (bid.vector[B + 2, ] - bid.vector[B + 1,
247 :                                       ])/(est.prob[B + 3, 3] - est.prob[B + 2,
248 :                                       3]) * (0 - est.prob[B + 3, 3])) +
249 :                                   bid.vector[B + 2, ] else trun.v
250 :                               est.prob[B + 4, 3] <- if(last.bid > trun.v) ((est.prob[B +
251 :                                       3, 3] - est.prob[B + 2, 3])/(bid.vector[B +
252 :                                       2, ] - bid.vector[B + 1, ])) * (trun.v -
253 :                                       bid.vector[B + 2, ])) + est.prob[B + 3,
254 :                                       3] else est.prob[B + 4, 3]
255 :                               last.bid <- if(last.bid > trun.v) trun.v else last.bid
256 :                               bids.nonp <- rbind(0, bid.vector, last.bid)
257 :                               est.prob.nonp <- est.prob[, 3]
258 :                               last.prob <- est.prob[B + 4, 3]

```

```

259 :      mont.nonp.mean <- nonpara.mean.wtp(bids.nonp,
260 :      est.prob.nonp)/(1 - last.prob)
261 :      mont.nonp.median <- nonpara.median.wtp(bids.nonp,
262 :      est.prob.nonp)
263 :      mont.mean[i, 1] <- mont.nonp.mean
264 :      mont.median[i, 1] <- mont.nonp.median
265 :      mont.est.prob[i, ] <- as.vector(est.prob[, 3])
266 :      ## parametric model ##
267 :      r.prob <- c(10, 0.5)
268 :      name <- c("CONS", "SIGM")
269 :      names(r.prob) <- name
270 :      mle.log <- nlminb(start = r.prob, obj = DBDCL.obj, X =
271 :      cv.datum)
272 :      mle.nor <- nlminb(start = r.prob, obj = DBDCP.obj, X =
273 :      cv.datum)
274 :      mle.wei <- nlminb(start = r.prob, obj = DBDCW.obj, X =
275 :      cv.datum)
276 :      est.coef.objec[i, (1:2)] <- mle.log$parameters
277 :      est.coef.objec[i, (3:4)] <- mle.nor$parameters
278 :      est.coef.objec[i, (5:6)] <- mle.wei$parameters
279 :      est.coef.objec[i, 7] <- mle.log$objective
280 :      est.coef.objec[i, 8] <- mle.nor$objective
281 :      est.coef.objec[i, 9] <- mle.wei$objective
282 :      coef.cons.l <- est.coef.objec[i, 1]
283 :      coef.sigm.l <- est.coef.objec[i, 2]
284 :      coef.cons.n <- est.coef.objec[i, 3]
285 :      coef.sigm.n <- est.coef.objec[i, 4]
286 :      coef.cons.w <- est.coef.objec[i, 5]
287 :      coef.sigm.w <- est.coef.objec[i, 6]
288 :      param.mean.l <- integrate(param.surv.l, 0.1, last.bid,
289 :      Xmax = last.bid, cons = coef.cons.l, sigm =
290 :      coef.sigm.l)$integral
291 :      param.mean.n <- integrate(param.surv.n, 0.1, last.bid,
292 :      Xmax = last.bid, cons = coef.cons.n, sigm =
293 :      coef.sigm.n)$integral
294 :      param.mean.w <- integrate(param.surv.w, 0.1, last.bid,
295 :      Xmax = last.bid, cons = coef.cons.w, sigm =
296 :      coef.sigm.w)$integral
297 :      param.median.l <- exp(coef.cons.l)
298 :      param.median.n <- exp(coef.cons.n)
299 :      param.median.w <- exp(coef.cons.w) * ((- log(0.5))^(
300 :      coef.sigm.w))
301 :      mont.mean[i, 2] <- param.mean.l
302 :      mont.mean[i, 3] <- param.mean.n
303 :      mont.mean[i, 4] <- param.mean.w
304 :      mont.median[i, 2] <- param.median.l
305 :      mont.median[i, 3] <- param.median.n
306 :      mont.median[i, 4] <- param.median.w
307 :    }
308 :    RESULT <- cbind(mont.mean, mont.median, est.coef.objec,
309 :    mont.est.prob)
310 :    names(RESULT) <- name.col
311 :    RESULT
312 :  }
313 :  ##### the function for summarizing a experiment result #####
314 :  result.dbdc.mont.exp.1 <- function(X, model, p.gamma, p.lambda)
315 :  {
316 :    ## X : a S data object made by function of 'dbdc.mont.exp.1'
317 :    ## model : model=1 means weibull model, model=2 means loglogistic model.
318 :    ## p.gamma, p.lambda : the parameters in the following density functions
319 :    ## ---weibull model
320 :    ## f(x)={p.gamma/p.lambda*(x/p.lambda)^(p.gamma-1)}*{exp-(x/p.lambda)^p.gamma}
321 :    ## ---loglogistic model

```

```

322 : ## f(x)={p.gamma*exp(-p.lambda)*x^(p.gamma-1)}/{1+exp(-p.lambda)*x^p.gamma}^2
323 :   gen.stat.data.set <- X
324 :   true.surv.l <- function(x, a, b)
325 :   {
326 :     1 - plogis( - b + a * log(x))
327 :   }
328 :   true.surv.w <- function(x, a, b)
329 :   {
330 :     1 - pweibull(x, a, b)
331 :   }
332 :   mean.logis <- p.lambda/p.gamma
333 :   true.mean <- if(model == 1) integrate(true.surv.w, 0.1, Inf, a =
334 :     p.gamma, b = p.lambda)$integral else integrate(
335 :     true.surv.l, 0.1, Inf, a = p.gamma, b = p.lambda)$
336 :     integral
337 :   true.median <- if(model == 1) p.lambda * log(2)^(1/p.gamma) else
338 :     exp(mean.logis)
339 :   result.matrix <- matrix(0, 10, 4)
340 :   result.matrix <- as.data.frame(result.matrix)
341 :   name.col <- c("Nonp", "logis", "norm", "Wei")
342 :   ## Nonp : nonparametric model
343 :   ## logis : loglogistic model
344 :   ## norm : lognormal model
345 :   ## Wei : weibull model
346 :   name.row <- c("mean.mean", "mean.var", "mean.s.d.", "mean.bias",
347 :     mean.mse-6, "median.mean", "median.var", "median.s.d.",
348 :     median.bias, "median.mse-6")
349 :   ## mean.mean : the mean of mean estimator
350 :   ## mean.var : the variance of mean estimator
351 :   ## mean.s.d. : the standard deviation of mean estimator
352 :   ## mean.bias : the bias of mean estimator
353 :   ## mean.mse-6: the mse*10^(-6) of mean estimator
354 :   ## median.mean : the mean of median estimator
355 :   ## median.var : the variance of median estimator
356 :   ## median.s.d. : the standard deviation of median estimator
357 :   ## median.bias : the bias of median estimator
358 :   ## median.mse-6: the mse*10^(-6) of median estimator
359 :   for(i in 1:4) {
360 :     result.matrix[1, i] <- mean(gen.stat.data.set[, i])
361 :     result.matrix[2, i] <- var(gen.stat.data.set[, i])
362 :     result.matrix[3, i] <- sqrt(var(gen.stat.data.set[, i]))
363 :     result.matrix[4, i] <- result.matrix[1, i] - true.mean
364 :     result.matrix[5, i] <- (result.matrix[4, i]^2 +
365 :       result.matrix[2, i]) * 10^(-6)
366 :     result.matrix[6, i] <- mean(gen.stat.data.set[, i + 4])
367 :     result.matrix[7, i] <- var(gen.stat.data.set[, i + 4])
368 :     result.matrix[8, i] <- sqrt(var(gen.stat.data.set[, i + 4]
369 :       )))
370 :     result.matrix[9, i] <- result.matrix[6, i] - true.median
371 :     result.matrix[10, i] <- (result.matrix[9, i]^2 +
372 :       result.matrix[7, i]) * 10^(-6)
373 :   }
374 :   dimnames(result.matrix) <- list(name.row, name.col)
375 :   result.matrix
376 : }

```

注

- 1) これは、農林水産省構造改善局計画部 [11] として一般に公開されている。
- 2) 二段階二肢選択形式は、ある金額を提示して、その金額を支払ってもよいかどうかを二度に渡って

質問する形式である。基本的には *yes*, *no* を回答する形式であるので、戦略的バイアスを引き起こしにくく、かつ一人の被験者の WTP が二肢選択形式よりも限定されるので、評価額の推定量を効率的に推定できることから、近年の我が国の CV 研究において、最も頻繁に採用される質問形式となっている。

- 3) パラメトリック推定法については、本補論 A を参照のこと。
- 4) この効果算定マニュアルは、農林水産省構造改善局計画部事業計画課 [13] として一般に公開されている。また、農業集落排水事業の効果算定マニュアルについても、農林水産省構造改善局計画部事業計画課 [12] として公開されている。
- 5) m^k/n^k が $\text{Prob}\{B^k \leq WTP\}$ の最尤推定量であることについては、対数尤度 $\ln L = m^k \ln \text{Prob}\{B^k \leq WTP\} + (n^k - m^k) \ln(1 - \text{Prob}\{B^k \leq WTP\})$ についての尤度方程式を解くことによって確認される。
- 6) Krström [6] もこれを採用している。
- 7) 例えば、McFadden [9]、栗山・北畠・大島 [7] などが挙げられる。
- 8) 但し、先に述べたように、ここでのケースにおいて、寺脇 [14] の推定量はそれぞれの提示額の生存確率についての最尤推定量となるが、全ての生存確率パラメータベクトルについての最尤推定量となるものではない。その意味では、この手法の非効率性が危惧されるが、筆者の試算によれば、両者の推定量の平均二乗誤差の差は極めて小さいという結果が得られている。このことについては稿を改めて論じたい。
- 9) パラメトリック推定法における最適提示額・配布数設計については、中嶋・寺脇 [10] を参照のこと。
- 10) 生存確率が 0 になる値は、最大提示額における生存確率とその一つ前の提示額における生存確率との間を結ぶ直線を外挿し、確率が 0 となるところの WTP を利用する。
- 11) MSE は推定量のバイアスと分散を統合した誤差の指標である。パラメータ τ の推定量 $\hat{\tau}$ の MSE は次のように表される。

$$\begin{aligned} \text{MSE}(\hat{\tau}) &:= E[(\hat{\tau} - \tau)^2] \\ &= E[\{\hat{\tau} - E(\hat{\tau})\}^2] + [E(\hat{\tau}) - \tau]^2 \\ &= \text{Var}(\hat{\tau}) + \text{Bias}(\hat{\tau})^2 \end{aligned}$$
- 12) 但し、ここでの結果は、プレテストなどで少なくとも中央値の位置がほぼ正確に把握されているケースに限られる。当然、中央値から遠く外れたところに少数の提示額を設計すれば、中央値が提示額間に含まれないことにより、誤差が大きくなることが予想され、その場合には、むしろ提示額を多く用意する方が、評価額の誤差がより小さくなることも十分に考えられる。
- 13) 地域用水環境整備事業の効果算定マニュアルとして公開されている農林水産省構造改善局計画部事業計画課 [13] では、提示額の数 7、配布部数は、回収率 50% を前提に 1,050 部とされており、農業集落排水事業の効果算定マニュアルとして公開されている農林水産省構造改善局計画部事業計画課 [12] では、提示額の数 5、配布部数は母集団全体とされている。
- 14) 以下のモデルの説明は、Hanemann, Loomis, and Kanninen [5]、および Alberini [1] を参考にした。
- 15) S-PLUS は Insightful Corporation (2001 年より、MathSoft Inc. から社名変更) の製品である。計算は、日本語版の Version 4.0J for Windows で行っている。日本での S-PLUS の入手方法については、(株)数理システムのホームページ <http://www.msi.co.jp/> を参照されたい。
- 16) ここで紹介する関数は、<http://www.taku-t.com/> からダウンロードすることができる。
- 17) 初期提示額数が M のとき、最大初期提示額のより高い二段階目の提示額、最小初期提示額のより低い二段階目の提示額を含めると、提示額数は合計 $M+2$ となる。

引用文献

- [1] Alberini, A., Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single-Bound, Double-Bound and Bivariate Models, *Journal of Environmental Economics and Management*, Vol. 28, 1995, pp. 287-306.
- [2] Ayer, M., H. D. Brunk, G. M. Ewing, and E. Silverman, An Empirical Distribution Function for Sampling with Incomplete Information, *Annals of Mathematical Statistics*, Vol. 26, 1955, pp. 641-647.
- [3] Carson, R. T., *Three Essays on Contingent Valuation*, Ph. D thesis, University of California, Berkeley, 1985.
- [4] Hanemann, W. M., Some Issues in Continuous- and Discrete-Response Contingent Valuation Studies, *Northeast Journal of Agricultural Economics*, Vol. 14, 1985, pp. 5-13.
- [5] Hanemann, W. M., J. B. Loomis, and B. Kanninen, Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation, *American Journal of Agricultural Economics*, Vol. 73, 1991, pp. 1255-1263.
- [6] Kriström, B., A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies, *Land Economics*, Vol. 66, 1990, pp. 135-139.
- [7] 栗山浩一・北畠能房・大島康行編著『世界遺産の経済学—屋久島の環境価値とその評価—』, 勁草書房, 2000.
- [8] McConnell, K. E., Models for Referendum Data: The Structure of Discrete Choice Models for Contingent Valuation, *Journal of Environmental Economics and Management*, Vol. 18, 1990, pp. 19-34.
- [9] McFadden, D., Contingent Valuation and Social Choice, *American Journal of Agricultural Economics*, Vol. 76, 1994, pp. 689-708.
- [10] 中嶋亮・寺脇拓「二段階二肢選択 CVM における最適な提示額設計」『農業経済研究』第70巻第4号, 1999, pp. 193-205.
- [11] 農林水産省構造改善局計画部監修『解説 土地改良の経済効果』大成出版社, 1997.
- [12] 農林水産省構造改善局計画部事業計画課『改訂版 農業集落排水事業における費用対効果分析マニュアル（案）』農林水産省, 2000.
- [13] 農林水産省構造改善局計画部事業計画課『水環境整備事業の効果算定マニュアル（案）』農林水産省, 2000.
- [14] 寺脇拓「都市近郊農業の外部経済効果の計測—二段階二肢選択 CVM における WTP 分布のノンパラメトリック推定—」『農業経済研究』, 第69巻第4号, 1998, pp. 201-212.
- [15] Turnbull, B. W., Nonparametric Estimation of Survivorship Function with Doubly Censored Data, *Journal of the American Statistical Association*, Vol. 69, 1974, pp. 169-173.