

Excel でできる CVM

第3. 2版

栗山 浩一（京都大学農学研究科）

2011年8月

〒606-8502

京都市左京区北白川追分町

京都大学農学研究科生物資源経済学専攻

TEL 075-753-6192

FAX 075-753-6191

E-mail: kkuri@kais.kyoto-u.ac.jp

*Excel は Microsoft 社の製品です

Excel でできる CVM

第3. 2版

栗山浩一(京都大学農学研究科)

要旨

生態系の価値を評価できる手法として仮想評価法(CVM)が注目を集めている。しかし、CVM で一般的に使われている二肢選択方式は統計的分析が必要であるため、これまでは統計専用のアプリケーションが必要であった。本研究は、初心者でも CVM を体験できるように表計算ソフトのみで推定する方法を示すとともに、推定方法の詳細を明らかにすることで統計ソフトがどのように計算しているかを示すことを目的としている。第3版ではフルモデルによる支払意志額の要因分析をサポートした。

キーワード： CVM、二肢選択方式、ロジット、ワイブル、最尤推定

1. はじめに

生態系などの非利用価値を評価できる数少ない手法として、近年、仮想評価法(contingent valuation method: CVM)が世界的に注目を集めている。CVM は環境の仮想的な変化を人々に示して、支払意志額(willingness to pay: WTP)や受入補償額(willingness to accept compensation: WTA)を回答者に直接たずねることで環境の経済価値を評価する手法である。

CVM では、表1に示されるように、数種類の質問形式が開発されている。初期の CVM 調査では自由回答形式や付値ゲームが用いられたが、これらの質問形式ではバイアスが生じやすいことが知られており、今日では二肢選択形式が一般的となっている。二肢選択形式とは、回答者にある金額を提示して、回答者に Yes/No のどちらかを答えてもらう方法である。Yes または No のどちらかを答えるだけなので、回答者にとっても答えやすく、またバイアスが比較的少ないことが知られている。

表 1 CVM の質問形式

自由回答形式(open-end) いくら払うかを自由回答でたずねる。無回答が多くなるという欠点がある。
付値ゲーム形式(bidding game) 例えば、最初に 1000 円以上払うかをたずねて、Yes ならば 2000 円、3000 円... と提示額を変化させて回答者の支払意志額に到達するまで繰り返し質問する。最初の提示額の影響を受ける可能性があること、および郵送法式では使えないという欠点がある。
支払カード形式(payment card) (1)100 円、(2)200 円、(3)500 円、... のように金額の選択肢を示して、どれかを選択してもらう。無回答が少なくなるものの、提示した金額の幅が回答に影響する可能性がある。
二肢選択形式(dichotomous choice) 例えば、1000 円以上払うかをたずねて、Yes/No のどちらかを1回だけたずねる。回答者が Yes と答える確率と提示額との関係から統計的に支払意志額を推定する。回答者が非常に答えやすく、信頼性の高い結果が得られるが、多数のサンプルを必要とする。

二肢選択形式には、一回だけ金額を提示するシングルバウンドと二回金額を提示するダブルバウンドがある。ダブルバウンドでは、最初に提示された金額に Yes と答えた場合はさらに高い金額を提示し、最初に No と答えた場合には低い金額を提示して二回たずねる。Hanemann, Loomis, and Kanninen (1991) が示したように、シングルバウンドよりもダブルバウンドの方が統計的効率性が高く、推定結果の信頼区間が狭まることを示している。ダブルバウンドは、少ないサンプルでも効率的に評価できるという利点がある¹。

ただし、二肢選択形式の Yes/No 回答から環境価値を推定するには統計分析が必要である。分析方法には、ランダム効用モデル、支払意志額関数モデル、生存分析などがあるが、いずれも統計分析専用のアプリケーションを必要とする。シングルバウンドについては既存の統計アプリケーションで推定できるものもあるが、ダブルバウンドについては多くの場合自分でプログラミングを行う必要がある²。このため、CVM に関心を持っているものの、統計アプリケーションを持っていないために評価できないケースや、統計アプリケーションを入手しても統計分析のプログラミング作業に苦しむケースがしばしば見られる。

そこで本研究では、統計アプリケーションを持っていないが CVM を体験してみたいという人や、統計アプリケーションがどのように推定を行っているのかを知りたい人を対象に、一般的に使われている表計算ソフト(EXCEL)を使って二肢選択形式の CVM による推定方法を紹介する。本研究で紹介するファイル(CVM.xls)を使えば、表計算ソフトを使える人ならば誰でも CVM による推定を体験することが可能である。

注意事項(必ずお読みください)

なお、このファイル(CVM.xls)はフリーウェアであるが、著作権は著者にある。著者に断りなく改変することは禁じる。本ファイル(CVM.xls)によって生じたいかなる損害も著者は一切の責任を負わない。また、本ファイル(CVM.xls)を用いて推定した結果等を論文、報告書、記事、ウェブサイト等で公表するときは、著者名や URL などを引用すること。引用を明記してあれば、著者に使用許可をとる必要はない。

2. CVM.xls の使い方(シングルバウンド)

CVM.xls は非常に簡単に使うことができる。ここでは実際の使い方について解説を行う。まず最初にシングルバウンド二肢選択形式の推定から説明する。ここではランダム効用モデルに基づいた対数線形ロジット分析による推定を行う。なお推定方法の詳細は後述する。

(1)ファイルを開いて「シングル ロジット」ワークシートを開く

¹ ダブルバウンドによる評価の実際例については栗山(1997)第3章および第4章を参照。

² CVM の評価に使われる代表的な統計アプリケーションについては栗山(1997)付録Bが紹介している。

図1 「シングル ロジット」のワークシート

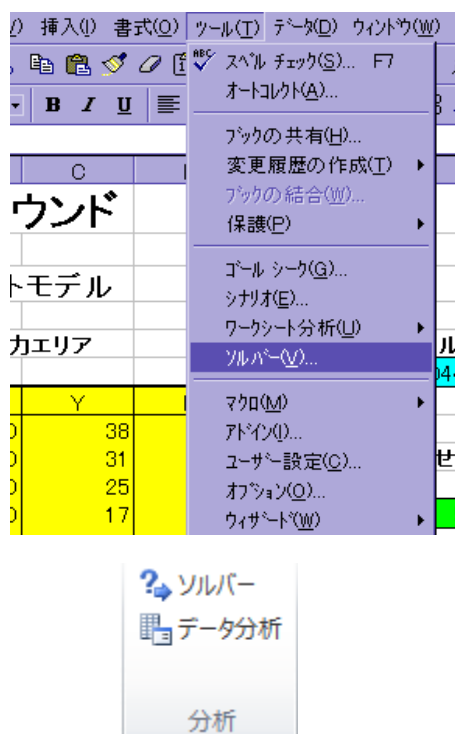
	A	B	C	D	E	F	G
1	シングルバウンド						
2							
3	対数線形ロジットモデル						
4							
5	データ入力エリア					目的セル	
6						-175.044	
7		Bid	Y	N			
8		500	38	8			
9		1000	31	12		変化させるセル	
10		2000	25	15		a	b
11		5000	17	23		5	-0.5
12		10000	17	28			
13		20000	8	36			
14							
15							

(2)データ入力エリアにデータを入力する

CVM のアンケート調査で得られたデータをデータ入力エリアに入力する。提示額を Bid の列に入力し、Yes 回答の数と No 回答の数をそれぞれ Y と N の列に入力する。一般的に低い提示額ほど Yes 回答が多く、提示額が高くなるにつれて No 回答が増える傾向が見られる。CVM がどのようなものを体験するだけならば、現在入力されているデータをそのまま使っても構わない。あるいは架空のデータを入力してみても構わない。提示額は 15 種類まで可能である。

(3)メニューから「ツール／ソルバー」を選択する

図2 ソルバーの選択



ツールの中に「ソルバー」の表示がない場合は「ソルバー」が組み込まれていないので、まず組み込む必要がある。ツールメニューから「アドイン」で「ソルバー」を選ぶとソルバーが組み込まれる。場合によってはExcelを再起動する必要があることもある。詳細は「ヘルプ」メニューから「ソルバーを組み込む」を検索する。

なお Excel 2007 以降では、「データ」タブに「ソルバー」があるので注意されたい。

(3)目的セルと変化させるセルを選択する

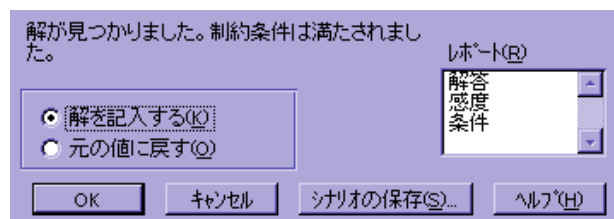
図3 目的セルと変化させるセルの選択



(4)実行ボタンを押すと推定が開始される

推定が完了すると図4が表示されるので、OK ボタンを押す。

図4 推定完了の画面



すると表2のように推定結果が表示される。表2は最初から入力されていたデータをそのまま分析したときの結果である。

表2 推定結果と推定 WTP

推定結果

変数	係数	t値	p値
constant	6.300399	6.694	0.000 ***
ln(Bid)	-0.76547	-6.680	0.000 ***
n			
対数尤度	-151.698		

推定WTP

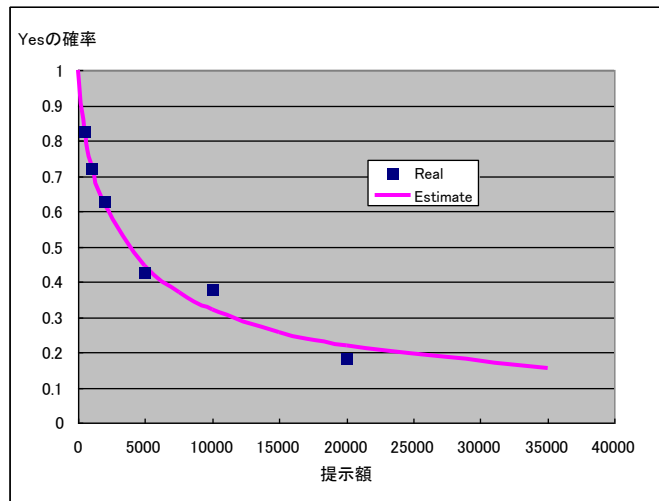
(中央値)	3,755
(平均値)	∞ 裾切りなし
	7,552 最大提示額で裾切り

変数は constant は定数項、ln(Bid)は提示額の対数値を意味する。係数を見ると、ln(Bid)の符号はマイナスなので提示額の対数値が高くなると回答者の効用が低下して Yes 回答の確率が低下することを示して

いる。t 値が高いので、どちらの変数も1%水準で有意となっている。有意水準は p 値が示している。***は1%水準、**は5%水準、*は10%水準で有意を意味する。

推定結果は図5のグラフで表示される。横軸は提示額、縦軸は Yes と回答する確率を示している。■は実際の回答結果、曲線は推定結果を意味する。図5を見ると推定結果は実際の回答結果にフィットしていることがわかる。

図5 推定結果のグラフ



支払意志額の中央値は Yes の確率が 0.5 となるところで、この場合は表2の通り 3,755 円である。平均値は曲線の下側の面積に相当する。無限大まで積分すると平均値は収束しないことがある。このため一般に最大提示額で積分計算を打ち切る。これを裾きりと呼ぶ。最大提示額で裾きりすると、表2の通り 7,552 円となる。同様に、「シングル ワイブル」のワークシートを使うとワイブルモデルによる推定も可能である。「シングル ノンパラ」のワークシートを使うと、 Kaplan-Meier モデルによる推定ができる。 Kaplan-Meier モデルの推定結果の解釈については、後述のターンブルモデルと同様なので参照されたい。

3. CVM.xls の使い方(ダブルバウンド)

次に二回金額を提示するダブルバウンドの場合について説明する。ダブルバウンドでは、ロジットモデル、ワイブルモデル、ターンブルモデルの3種類が用意されている。まずロジットモデルから説明する。

図6 ダブルバウンド(ロジット)のワークシート

	A	B	C	D	E	F	G	H
1	ダブルバウンド							
2								
3	対数線形ロジットモデル							
4								
5	データ入力エリア							
6								
7		T1	TU	TL	YY	YN	NY	NN
8	1	1000	3000	500	18	25	3	23
9	2	3000	6000	1000	10	19	13	34
10	3	6000	15000	3000	6	14	8	49
11	4	15000	40000	6000	2	18	5	49
12	5							
13	6							
14	7							
15	8							
16	9							
17	10							
18								
19								
20		目的セル		変化させるセル				
21		-354.40576		a	b			推定結果はこちら →
22				6.68875385	-0.9093837			

シングルバウンドとの使い方の違いはデータの入力部分のみである。ダブルバウンドの場合は、一回目の提示額に Yes と回答した場合は高い金額を提示し、No と回答した場合には低い金額を提示する。そこで、一回目の提示額を T1 の列、二回目の高い提示額を TU の列、二回目の低い提示額を TL の列に入力する。

回答結果は二回とも Yes の回答数を YY の列、一回目は Yes で二回目は No の回答数を YN の列、一回目は No で二回目は Yes の回答数を NY の列、そして二回とも No の回答数を NN の列に入力する。提示額の種類は 10 種類まで可能である。

後は、シングルバウンドと同様にソルバーで目的セルと変化させるセルを選択すると推定が開始される。推定結果は表3のように表示される。推定結果や支払意志額の読み方はシングルバウンドと全く同じである。

表3 ダブルバウンド ロジット推定結果

推定結果

変数	係数	t値	p値
constant	6.6888	10.631	0.000***
ln(Bid)	-0.9094	-11.304	0.000***
n	296		
対数尤度	-354.406		

推定 WTP

(中央値)	1,564
-------	-------

(平均値)	∞ 裾切りなし 5,751 最大提示額で裾切り
-------	-----------------------------------

二番目のモデルはワイブルモデルである。ワイブル分布は非常に柔軟な関数なので、一般にワイブルモデルの当てはまりは非常によい。ワイブルモデルの計算方法はロジットモデルと全く同じである。ワイブルモデルの推定結果は表4のとおり。対数尤度がロジットよりも高く、ワイブルの当てはまりがよいことを示している。また、ワイブルモデルでは、平均値の計算で無限大まで積分しても発散しないことが多く、裾きりの影響が比較的少ない。

表4 ダブルバウンド ワイブル推定結果

変数	係数	t値	p値
Location	8.0482	60.842	0.000***
Scale	1.8878	12.430	0.000***
n	296		
対数尤度	-350.414		

推定 WTP

(中央値)	1,566
-------	-------

(平均値)	5,655 裾切りなし 5,173 最大提示額で裾切り
-------	--------------------------------

三番目のモデルはターンブルモデルである。ターンブルは、これまでのモデルとは異なり、分布関数を仮定しないノンパラメトリックな推定方法である。ターンブルモデルによって推定するには「ダブルバウンド ノンパラ」のワークシートを用いる。

ターンブルモデルの場合は、データ入力で注意すべき点がある。提示額の設計に注意が必要である。例えば、第二バージョンの最初の提示額 T1 は 3000 円だが、TU の 6000 円は第三バージョンの T1 と等しくする。同様に第二バージョンの TL の 1000 円は第一バージョンの T1 と等しくする。他のバージョンについても同様である。

図7 「ダブルバウンド ノンパラ」ワークシート

	A	B	C	D	E	F	G	H
1	ダブルバウンド							
2								
3	ノンパラメトリック生存分析(Turnbull)							
4								
5								
6								
7								
8	データ入力エリア							
9								
10		T1	TU	TL	YY	YN	NY	NN
11	1	1,000	3,000	500	18	25	3	23
12	2	3,000	6,000	1,000	10	19	13	34
13	3	6,000	15,000	3,000	6	14	8	49
14	4	15,000	40,000	6,000	2	18	5	49
15	5							

使い方はこれまでと同様に、ソルバーで目的セルと変化させるセルを選択すると推定が開始される。なお、図8のように「初期値」が与えられているので、これを参考に变化させるセルに値を入れるとよい。推定結果は表5のとおり。Lower は支払意志額の下限、Upper は上限である。生存確率は支払意志額が Upper 以上である確率を意味する。例えば、支払意志額が 500 円以上の確率は 0.658 となる。金額が高くなるにつれて生存確率は低下する。

図8 ターンブルモデルの設定

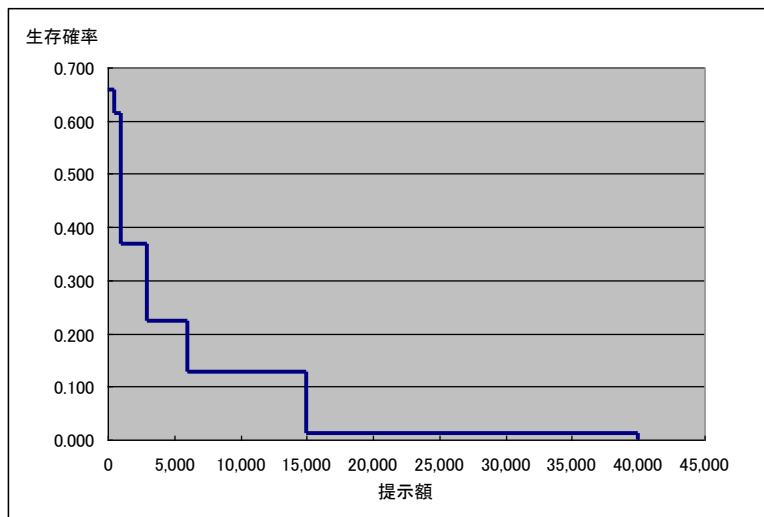
目的セル	変化させるセル	初期値
-390.564	1	0.1429
	2	0.1429
	3	0.1429
	4	0.1429
	5	0.1429
	6	0.1429
	7	0.1429
	8	0.1429
	9	0.1429
	10	0.1429
	11	0.1429
	12	0.1429

表5 Turnbull 推定結果

推定結果

Lower	Upper	生存確率	t値	P値
0	500	0.658	16.44	0.0000 ***
500	1000	0.614	17.01	0.0000 ***
1000	3000	0.369	12.31	0.0000 ***
3000	6000	0.222	8.76	0.0000 ***
6000	15000	0.129	5.85	0.0000 ***
15000	40000	0.013	1.44	0.1497
40000	+∞	0.000		
n	296			
対数尤度	-344.36			

図9 生存曲線のグラフ



またこれまでのモデルと同様に推定結果がグラフで表示される。グラフの見方はシングルバウンドのときと同様で、横軸は提示額、縦軸は生存確率 (YES の確率) である。シングルバウンドのときと異なり、ターブルはノンパラメトリック法なので推定結果は階段状に表示される。支払意志額の中央値は生存曲線が 0.5 となるところであり、表6のように区間で表示される。この場合は中央値は 1000~3000 円となる。これは、中央値がこの区間内にあることを意味している。いわゆる信頼区間とは異なるので注意されたい。平均値は下限値と中位平均値が示されている。

表6 ダブルバウンドの推定 WTP

推定WTP

中央値

1000	~	3000
------	---	------

平均値

3,520
5,949

4. フルモデル

支払意志額には、年齢や所得などの個人属性や評価対象に対する知識などの要因が影響する可能性がある。こうした支払意志額の要因を分析するのがフルモデルである。フルモデルでは、個人別のデータを「フルモデル用データ」のワークシートに入力してから分析を行う(図 10)。

ID は各回答者の ID 番号である。一回だけ金額を提示するシングルバウンドの場合は「シングルバウンド」と書かれた部分にデータを入力し、「ダブルバウンド」と書かれた部分はゼロを入力する。二回金額を提示するダブルバウンドの場合は、「シングルバウンド」と「ダブルバウンド」の両方の部分に入力する。

図 10 「フルモデル用データ」ワークシート

	A	B	C	D	E	F	G	H
1	フルモデル用データ入力エリア							
2								
3		入力方法						
4			各回答者別にデータを入力。					
5			シングルバウンドのときはY, N, T1にデータを入力、YY, YN, NY, NI					
6			ダブルバウンドのときはYY, YN, NY, NN, T1, TU, TLにデータを入力					
7			説明変数はx1~x10に入力。					
8			サンプル数が500を超えるときは一番下側の行をコピーして必要					
9								
10								
11			シングルバウンド用			ダブルバウンド用		
12								
13		ID	Y	N	YY	YN	NY	NN
14		1	1	0	0	1	0	0
15		2	0	1	0	0	0	1
16		3	0	1	0	0	1	0
17		4	0	1	0	0	0	1
18		5	0	1	0	0	1	0
19		6	1	0	1	0	0	0

たとえば、図の ID が 1 番の回答者の場合、最初の提示額に対しては「Yes」、二回目の提示額については「No」と回答している。このとき、最初の提示額のみ用いるシングルバウンドでは、最初の提示額に対して Yes なので「Y」のデータは「1」、「N」のデータは「0」となる。一方、ダブルバウンドのときは最初は「Yes」、二回目は「No」と回答しているので「YN」の部分で「1」とし、それ以外の「YY」「NY」「NN」は「0」となる。

図 11 提示額および説明変数

シングルバウンド用						
ダブルバウンド用			ここから説明変数。			
T1	TU	TL	x1	x2	x3	
2000	5000	1000	1	1	0	
2000	5000	1000	0	0	1	
1000	2000	500	1	0	0	
1000	2000	500	1	0	1	
2000	5000	1000	1	1	1	
5000	10000	2000	1	0	0	

提示額については、図 11 のように最初の提示額「T1」、Yes 回答時の提示額「TU」、No 回答時の提示額「TL」を入力する。シングルバウンドのときは「T1」のみ入力し、「TU」と「TL」はゼロを入力する。例えば、ID 番号 1 番の回答者には最初に 2000 円が提示され、Yes 回答時には 5000 円、No 回答時には 1000 円が提示されることを示している。

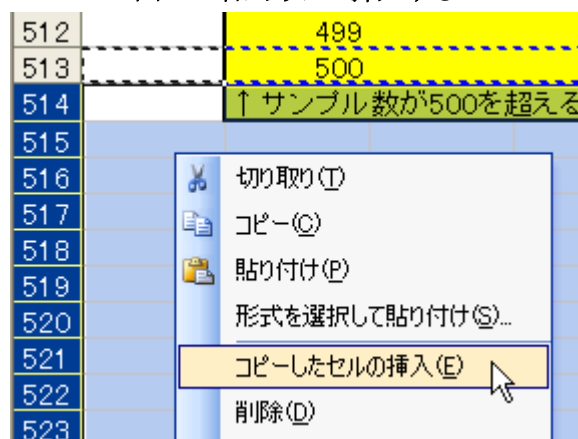
支払意志額の説明要因は説明変数の「x1」～「x10」の部分に入力する。説明変数は性別のように 0 または 1 の値のみとるダミー変数でもかまわないし、年齢や所得のような数値変数でもかまわない。使用しない説明変数の場所は「0」を入力しておく。例えば「x1」～「x5」しか使用しないときは「x6」～「x10」はすべて「0」を入力する。

データは現在は 500 人のデータまで入力できるようになっている。500 人を超える場合は、図 12 のように ID 番号 500 の行をコピーし、マウスの右クリックの「コピーしたセルの挿入」で必要な人数分だけ挿入してから、データを入力する。「貼り付け」ではなく「挿入」なので注意すること。

図 12 回答者が 500 人を超える場合

511	498
512	499
513	500
514	↑ サンプル数が500を超えるときはこの上の行(ID 500)を:
515	

図 13 右クリックで挿入する



データが入力できたらシングルバウンドの場合は「フルモデル シングル」、ダブルバウンドの場合は「フルモデル ダブル」のワークシートで推定を行う。図 14 はダブルバウンド用のワークシート画面である。説明変数のうち、モデルに入れるものは 1、モデルに入れないものは 0 を設定する。この図の場合、「x3」、「x5」、「x7」は支払意志額の説明要因として使われない。

「ツール」メニュー→「ソルバー」で、目的セルと変化させるセルを設定し、「実行」をクリックすると推定が行われる。計算にはしばらく時間がかかるが、ワークシート左下に計算の施行状況が表示される。

図 14 フルモデル推定用ワークシート

フルモデル ダブルバウンド

対数線形ロジットモデル

目的セル	-1486.59	収束していません
変化させるセル		モデルに入れる=1
定数	2	↓
ln(Bid)	-1	↓
x1	0	1
x2	0	1
x3	0	0
x4	0	1
x5	0	0
x6	0	1
x7	0	0
x8	0	1
x9	0	1
x10	0	1

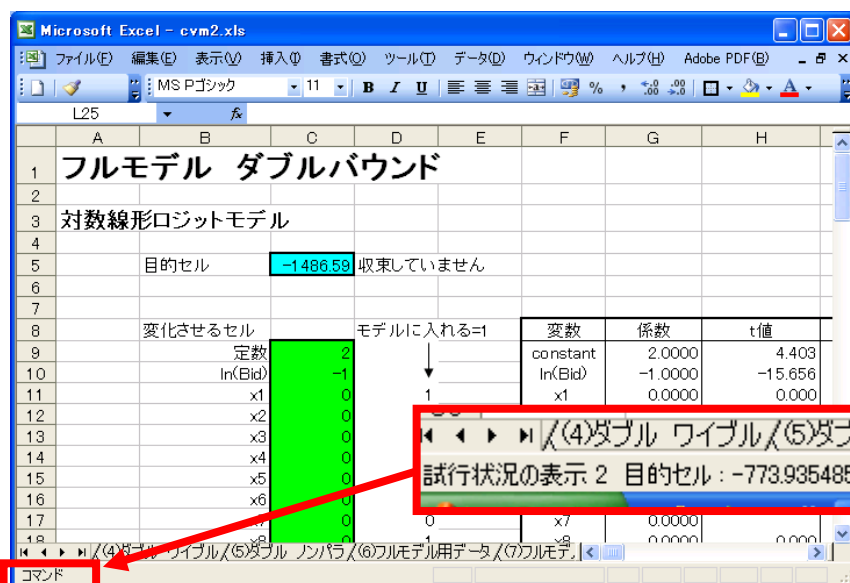


表7 フルモデル推定結果

変数	係数	t値	p値
constant	8.3485	11.493	0.000 ***
ln(Bid)	-1.4278	-14.488	0.000 ***
x1	0.0503	0.256	0.798
x2	1.1497	5.565	0.000 ***
x3	0.0000		
x4	-0.5381	-2.668	0.008 ***
x5	0.0000		
x6	0.5426	7.179	0.000 ***
x7	0.0000		
x8	-0.1168	-1.672	0.095 *
x9	0.1249	1.740	0.083 *
x10	0.0018	6.046	0.000 ***
n	400		
対数尤度	-475.5254		

しばらくすると表7のような推定結果が表示される。係数の符号がプラスのものは支払意志額にプラスの影響を及ぼし、逆に符号がマイナスのものは支払意志額にマイナスの影響を与える。説明変数として用いられなかったものは係数が「0.0000」となっている。各説明変数の有意水準をp値で判断し、有意ではない変数はモデルから削除する。たとえば、表7の場合、「x1」のp値は 0.798 であり「*」が一つもないので10%水準でも有意ではない。したがって、「x1」は支払意志額には影響していないと考えられるので、「モデルに入れる」の部分で「0」にして「x1」をモデルから削除して、再度、推定を行う。このプロセスを繰り返すことで、最終的なモデルを決定する。

5. 推定方法

ここでは CVM.xls の推定方法の詳細について説明する。これ以降は統計学の知識を必要とするため、初心者の方は読み飛ばしても構わない。

5.1 シングルバウンド ロジット

シングルバウンド・ロジットでは Hanemann (1984)のランダム効用モデルを用いている。仮想的な環境政策を回答者に提示し、環境政策が実施されたときには負担額が T 円だけかかるとする。一方、環境政策が実施されない場合の負担額は0円とする。このとき、負担額T円で環境政策が実施されたときの効用関数を U_Y 、負担額0円で政策が実施されないときの効用関数を U_N とする。効用関数は観察可能な V_Y 、 V_N と誤差項 ε_Y 、 ε_N によって構成されるとする。負担額T円の環境政策に対して回答者がYESと答える確率は、環境政策が実施されたときの効用が実施されないときの効用よりも高い確率であるから、

$$\begin{aligned}
 \Pr[\text{Yes}] &= \Pr[U_Y > U_N] \\
 &= \Pr[V_Y + \varepsilon_Y > V_N + \varepsilon_N] \\
 &= \Pr[\varepsilon > -\Delta V]
 \end{aligned}
 \tag{1}$$

ただし、 $\Delta V = V_Y - V_N = \beta' \mathbf{x}$ 、 $\varepsilon = \varepsilon_Y - \varepsilon_N$ である。 β は推定されるパラメータのベクトル、 \mathbf{x} は説明変数ベクトルである。ここでは、効用差関数として対数線形関数 $\Delta V = \beta_0 + \beta_T \ln T + \sum \beta_k x_k$ を用いている。

ここで、 ε_Y および ε_N が第一種極値分布(Gumbel 分布)に従うと仮定すると ε がロジスティック分布となりロジットモデルが適用できる。ロジットモデルでは、回答者が YES と答える確率は、以下の通りとなる。

$$\Pr[\text{Yes}] = \frac{1}{1 + \exp(-\Delta V)} = \Lambda \quad (2)$$

このとき対数尤度関数は

$$\ln L = \sum_i [d_Y \ln \Pr[\text{Yes}] + d_N \ln(1 - \Pr[\text{Yes}])] \quad (3)$$

となる。ただし、 d_Y は回答者が YES と答えたときに1となるダミー変数、 d_N は NO と答えたときに1となるダミー変数である。パラメータの推定は最尤法により行われる。つまり(3)式が最大となるようにパラメータ β が推定される。(3)式の一階の条件は

$$\frac{\partial \ln L}{\partial \beta} = \sum_i [d_Y - \Lambda] \mathbf{x} = \mathbf{0} \quad (4)$$

となる。したがって(4)を解くことによりパラメータの推定が行われる。また(3)式の二階の条件は

$$\mathbf{H} = \frac{\partial^2 \ln L}{\partial \beta \partial \beta} = -\sum_i \Lambda(1 - \Lambda) \mathbf{x} \mathbf{x}' \quad (5)$$

(5)式より対数尤度関数は大域的に凹関数であり、最大点は唯一であることを示している。

最尤法により推定されたパラメータをもとに支払意志額を算出する。支払意志額には中央値と平均値の2種類がある。中央値は YES と答える確率が 0.5 となるときの提示額に相当する。ロジットモデルのときは(2)式より効用差 ΔV が0となるときに相当する。したがって、効用差関数が対数線形の場合、支払意志額の中央値は次式により算出される。

$$\text{中央値 } WTP^* = \exp\left(-\frac{\beta_0 + \sum \beta_k x_k}{\beta_T}\right) \quad (6)$$

一方の平均値の場合は、 $\int_0^\infty \Lambda(T) dT$ によって算出される。効用差関数が対数線形の場合は、

Hanemann (1984)が示したように支払意志額の平均値は次の通りとなる。

$$\text{平均値 } WTP^+ = -\exp\left(-\frac{\beta_0 + \sum \beta_k x_k}{\beta_T}\right) \frac{\pi / \beta_T}{\sin(-\pi / \beta_T)}, \text{ ただし } 0 > \frac{1}{\beta_2} > -1 \quad (7)$$

最大提示額で裾きりするときは、積分計算を 0 から最大提示額まで行う。これは数値計算により行う。

5.2 シングルバウンド ワイブル

ワイブルモデルは生存分析を用いている。生存分析では生物統計学や経営工学で使われる統計手法である。提示額 T のときに Yes と回答する確率を示す関数を生存関数 $S(T)$ と呼ぶ。なお生存関数 $S(T)$ と分布関数 $G(T)$ には $S(T) = 1 - G(T)$ の関係がある。加速ワイブルモデルでは、生存関数として次のようなワイブル分布関数を想定する。

$$S(T) = \exp\left(-\exp\left(\frac{\ln T - \mu}{\sigma}\right)\right)$$

μ は位置パラメータ、 σ はスケールパラメータと呼ばれている。この場合、支払意志額の平均値と中央値は、以下のとおりとなる。ただし、 Γ はガンマ関数である。

$$\text{平均値 } \exp(\mu) \cdot \Gamma[1 + \sigma]$$

$$\text{中央値 } \exp(\mu) \cdot [-\ln(0.5)]^\sigma$$

5.3 シングルバウンド ノンパラメトリック

シングルバウンドのノンパラメトリックでは、カプラン・マイヤーモデルが使われる。提示額 T_j 円に対して Y_j 人が Yes を選択し、 N_j 人が No を選択したとする。提示額 T_j 円に対して回答者が Yes と答える確率を p_j とする。このとき対数尤度関数は以下の通りとなる。

$$\ln L = \sum_j Y_j \ln p_j + N_j \ln(1 - p_j)$$

対数尤度関数の一階の条件は以下のとおりとなる。

$$\frac{\partial \ln L}{\partial p_j} = \frac{Y_j}{p_j} - \frac{N_j}{1 - p_j} = 0$$

これを解くと以下が得られる。つまり、 T_j が提示された人数のうち Yes を選択した人数の比率が生存確率の最尤推定量となる。

$$p_j = \frac{Y_j}{Y_j + N_j}$$

また二階の条件は以下のとおりである。

$$\frac{\partial^2 \ln L}{\partial p_j \partial p_j} = -\frac{Y_j}{p_j^2} - \frac{N_j}{(1-p_j)^2}$$

したがって、対数尤度関数は大域的に凹関数であり、最大点は唯一となる。

5.4 ダブルバウンド

ダブルバウンドではロジット、ワイブル、ターンプルの3種類が用意されている。ロジットはシングルバウンドと同じランダム効用モデルを用いている。ワイブルとターンプルは生存分析である。ダブルバウンドでは回答は、YY, YN, NY, NN の4種類が得られる。このとき、それぞれの回答が得られる確率は、

$$\Pr[YY] = 1 - G(TU) = S(TU)$$

$$\Pr[YN] = G(TU) - G(T1) = S(T1) - S(TU)$$

$$\Pr[NY] = G(T1) - G(TL) = S(TL) - S(T1)$$

$$\Pr[NN] = G(TL) = 1 - S(TL)$$

となる。ただし、G(T)は提示額がTのときの分布関数、S(T)は生存関数である。ロジットの場合は次の分布関数を用いる。

$$\text{ロジット } G(T) = \frac{1}{1 + \exp(\beta_0 + \beta_T \ln T + \sum \beta_k x_k)}$$

ワイブルの場合は次の生存関数を用いる

$$S(T) = \exp\left(-\exp\left(\frac{\ln T - \mu}{\sigma}\right)\right)$$

ノンパラメトリックでは Turnbull (1976)のノンパラメトリック区間打ち切り生存分析を用いた。ダブルバウンドでは二回たずねているので、WTPはある区間に入ることがわかる。例えば、1000 円では Yes、3000 円では No のときは WTP が 1000~3000 円の区間の中にある。1000 円で Yes、3000 円でも Yes のときは WTP は 3000 円~+∞の区間にあり、1000 円で No、500 円でも No のときは WTP は 0~500 円の区間の中にある。

提示額 T_j 円に対して回答者が Yes と答える確率を p_j とする。この確率は生存確率と呼ばれる。WTP が $T_j \sim T_{j+1}$ の区間にあるとすると、対数尤度関数は以下の通りとなる。

$$\ln L = \sum_j N_j \ln(p_j - p_{j+1}) \quad (8)$$

となる。ただし、 N_j は WTP が $T_j \sim T_{j+1}$ の区間にある回答者の人数である。

Turnbull では生存確率が直接推定される。生存確率は(8)式が最大となるように推定される。(8)式の一階の条件は

$$\frac{\partial \ln L}{\partial p_j} = \frac{N_j}{p_j - p_{j+1}} - \frac{N_{j-1}}{p_{j-1} - p_j} = 0 \quad (9)$$

となる。生存確率 p_j は(9)式を解くことで得られる。また(8)式の一階の条件は

$$\frac{\partial^2 \ln L}{\partial p_j \partial p_j} = -\frac{N_{j-1}}{(p_{j-1} - p_j)^2} - \frac{N_j}{(p_j - p_{j+1})^2}, \quad \frac{\partial^2 \ln L}{\partial p_j \partial p_{j+1}} = \frac{N_{j-1}}{(p_j - p_{j+1})^2} \quad (10)$$

となり、対数尤度関数は大域的に凹関数であり、最大点は唯一となる。

支払意志額の算出には、シングルバウンドと同様に中央値と平均値がある。中央値の場合は、生存確率が 0.5 となるところであるが、Turnbull ノンパラメトリック法では生存確率は階段状の曲線となるため中央値は 0.5 を含む区間として推定され、点推定はできない。一方の平均値については、次式により下限値を求めることができる。

$$E[WTP]_{LOWER} = \sum_j T_{j-1}(p_j - p_{j-1}) \leq \int_0^\infty T dp(T) = E[WTP]$$

5.5 最尤法

最尤法は、対数尤度関数が最大となるようにパラメータの推定を行う。一般に対数尤度関数は非線型なので、最適解を直ちに求めることが困難であり、試行錯誤により最適解を求める必要がある。繰り返し回数 t 回目のときのパラメータを θ_t とすると、次式によりパラメータの更新を行う。

$$\theta_{t+1} = \theta_t + \lambda_t \Delta_t \quad (11)$$

ただし、 λ はステップサイズ、 Δ は方向ベクトルである。このパラメータの更新方法のアルゴリズムには、Newton 法、BHHH、DFP、BFGS など様々な方法が開発されているが、ここでは最も簡単な Newton 法について解説する。対数尤度関数の一階の条件を θ_t でテーラー展開すると、

$$\frac{\partial \ln L(\theta_{t+1})}{\partial \theta_{t+1}} = \frac{\partial \ln L(\theta_t)}{\partial \theta_t} + H(\theta_t)(\theta_{t+1} - \theta_t) = 0 \quad (12)$$

ただし、 H は対数尤度関数のヘッセ行列である。

$$H(\theta_t) = \frac{\partial^2 \ln L(\theta_t)}{\partial \theta_t \partial \theta_t} \quad (13)$$

(12)式を θ_{t+1} について解くと次式が得られる。

$$\theta_{t+1} = \theta_t + (H(\theta_t))^{-1} \frac{\partial \ln L(\theta_t)}{\partial \theta_t} \quad (14)$$

ロジットの場合はヘッセ行列が(5)式のように簡単な形となるので、このニュートン法により最尤法の最適解が得られる。

最尤法の推定値は漸近的に正規となる性質がある。つまり、 $\hat{\theta}_{ML} \rightarrow N[\theta, \{I(\theta)\}^{-1}]$ が成立する。

ただし、 N は正規分布、 $I(\theta) = -E[\partial^2 \ln L / \partial \theta \partial \theta']$ はフィッシャー情報行列である。これより、パラメータの推定値の標準誤差が得られ、推定値の検定が可能となる。

なお、ヘッセ行列を厳密に計算することが容易ではない場合は、BHHH 推定量が用いられる。これは

$$I(\theta) = \sum_{i=1}^n \hat{\mathbf{g}}_i \hat{\mathbf{g}}_i' = [\hat{\mathbf{G}}_i \hat{\mathbf{G}}_i']$$

ただし、

$$\hat{\mathbf{g}}_i = \frac{\partial \ln L(\theta)}{\partial \theta}, \quad \hat{\mathbf{G}} = [\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \dots, \hat{\mathbf{g}}_n]$$

である。CVM.xls では、ダブルバウンドのロジットおよびワイブルではヘッセ行列の計算に BHHH を用いている。

6. おわりに

本研究では、統計アプリケーションを用いずに、表計算ソフト(EXCEL)のみで二肢選択形式のデータを分析する方法を紹介した。二肢選択形式はバイアスが比較的少なく、今日ではもっとも標準的な質問形式となっているが、統計分析が複雑なため初心者には推定が難しいという問題があった。このため、CVM を実際に試してみたいと思っても、統計分析の段階で壁にぶつかるケースがしばしば見られる。このような場合、表計算ソフトのみで二肢選択形式を体験できることは、初心者にとって有用なであろう。

また統計アプリケーションを用いるにしても、現段階ではシングルバウンドについては比較的簡単に推定できるものの、ダブルバウンドについては自分でプログラムを組むことが必要となる場合がほとんどである。そのため、本格的に CVM を研究するならば、統計アプリケーションをブラックボックスとして使うのではなく、CVM の推定方法について詳細を知っておくことは重要である。このような場合も、CVM.xls のワークシートを参照することで推定方法の詳細を知ることができるであろう。

このように、本研究で紹介した CVM.xls を用いることで初心者であっても CVM を体験できるようになる。ただし、実際に CVM で環境価値を評価するときには、統計分析以外にも注意しなければならない点が多数あることを忘れてはならない。慎重にサーベイデザインを行わなければバイアスが生じる危険性があるし、二肢選択形式の場合は提示額設計が非常に重要となる。そしてサーベイデザインや提示額設計の問題点を確認するためには、プレテストを繰り返す必要がある。

本研究は、CVM の統計分析を体験したり理解することに役立つであろうが、CVM は統計分析だけで結果が出るものではない。なぜなら統計分析は一連の CVM の評価プロセスの一つにすぎないからである。サーベイデザイン、提示額設計、プレテスト、統計分析、評価結果の信頼性の確認など、様々な評価プロセスのすべてが CVM の評価には不可欠であることに注意が必要である。

参考文献

二肢選択形式の推定方法については

Hanemann, W. M. (1984) Welfare Evaluations in Contingent Valuation Experiments with Discrete Response, *American Journal of Agricultural Economics*, 66(3): 332-341

ダブルバウンドの推定方法については

Hanemann, M., Loomis, J., and Kanninen, B. (1991) Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics*, 73(4): 1255-1263.

生存分析については

Nelson, W. (1982) *Applied Life Data Analysis*. New York: Wiley.

最尤法については

Greene, W. H. (2003) *Econometric Analysis*, 5th ed. Prentice Hall.

CVMをわかりやすく解説したものは

栗山浩一(1997)「公共事業と環境の価値－CVMガイドブック－」築地書館

栗山浩一(2000)「図解 環境評価と環境会計」日本評論社

環境評価全般をわかりやすく解説したものは

鷲田豊明(1999)「環境評価入門」勁草書房

環境評価手法の最近の動向を解説したものは

鷲田豊明・栗山浩一・竹内憲司編(1999)「環境評価ワークショップ－評価手法の現状」築地書館

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

CVMの政策利用については

竹内憲司(1999)「環境評価の政策利用」勁草書房

CVMの理論を詳しく解説したものは

栗山浩一(1998)「環境の価値と評価手法－CVMによる経済評価－」、北海道大学図書刊行会

土木分野での評価については

森杉壽芳編(1997)「社会資本整備の便益評価：一般均衡理論によるアプローチ」勁草書房

肥田野登(1997)「環境と社会資本の経済評価：ヘドニック・アプローチの理論と実際」勁草書房

大野栄治編(2000)「環境評価の実務」勁草書房

農業経済分野での評価については

出村克彦・吉田謙太郎編(1999)「農村アメニティの創造に向けて－農業・農村の公益的機能評価」大明堂

浅野耕太(1998)「農林業と環境評価」多賀出版