

第 11 章 世の中白と黒じゃない(修正版)

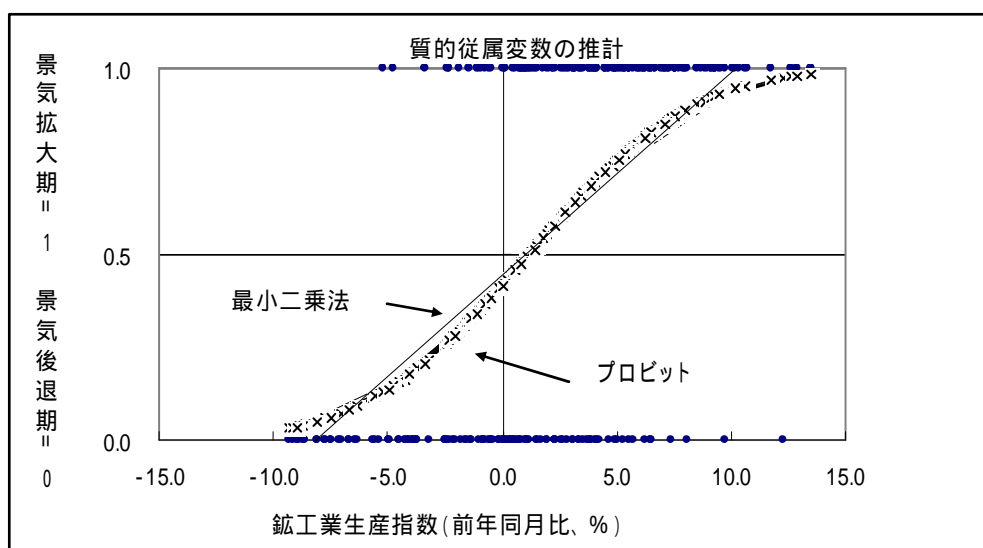
(1) 質的従属変数の推計

質的従属変数とは、被説明変数が 1 かゼロである場合の推計法である。「働くか働かないか」、「自動車を持っているか持っていないか」などを 1 とゼロのデータとして表して推計に使う。

下の図は、被説明変数(従属変数)に景気の状態をとったものであり、景気拡大期は 1、景気後退期は 0 とする。1 と 0 上にある各点はそれに対応する鉱工業生産指数の前年同月比の値である。

こうしたデータに最小二乗法を適用すると、推計値がゼロを下回ったり、1 を上回っている場合が出てくる。これを回避するために、ある曲線をゼロと 1 の間にあてはめようというのがプロビット、ロジットの考え方である。

プロビットは累積正規分布曲線上に当てはめるようにし、ロジットはロジスティック曲線上に当てはめるようにする。

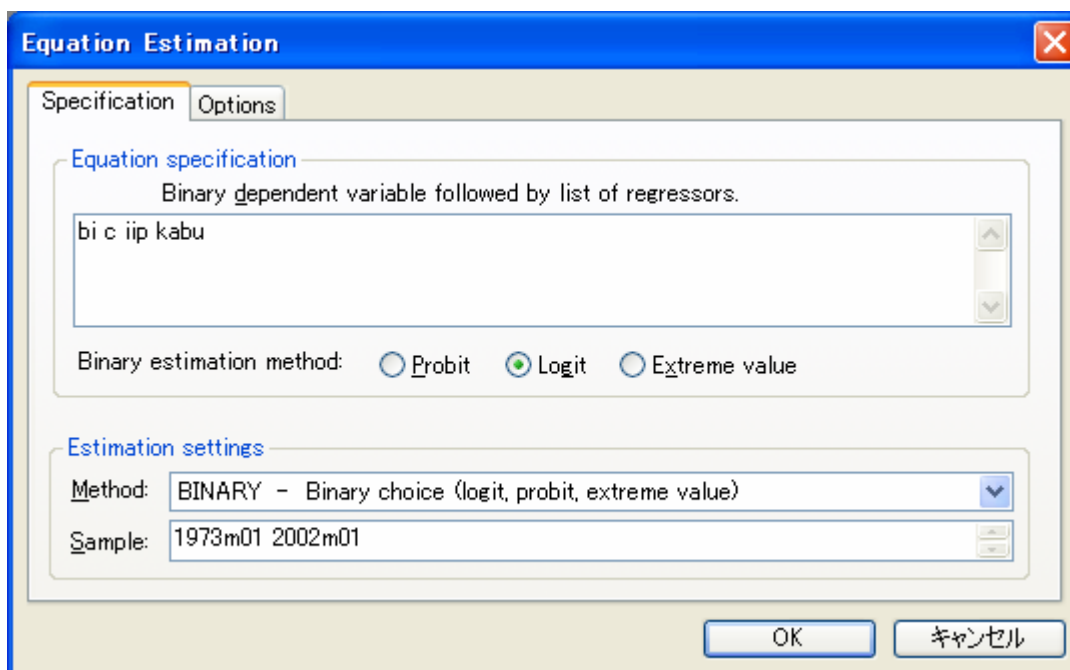


(2) EViews での推計法

最小二乗法など推計法を選ぶメニューで、「BINARY」を選ぶ。

Quick Estimate Equation...

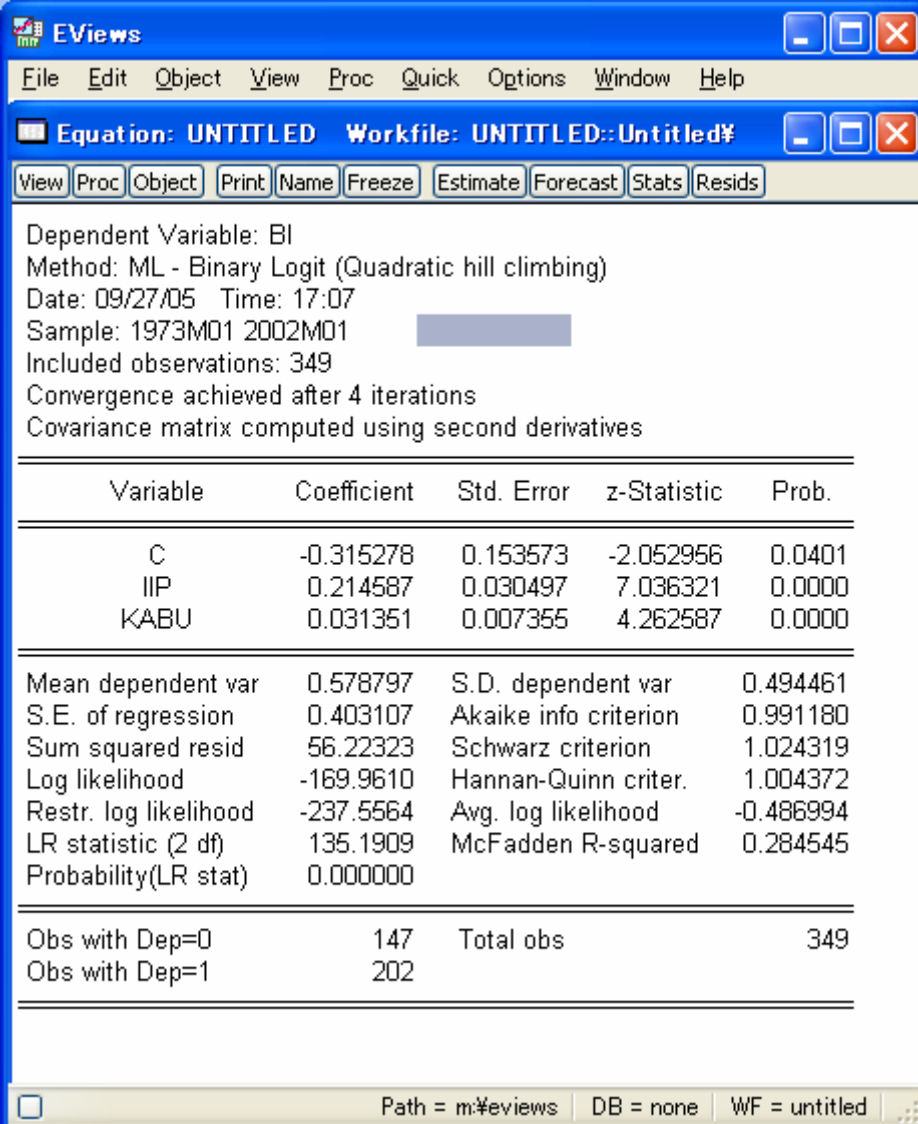
Method を BINARY に変更する。



質的従属変数の推計法で代表的なものは、プロビットとロジットである。プロビットの場合は、binary estimation method を Probit に、ロジットの場合は Logit にする。

(3) 推計結果の解釈

1973年1月から2002年1月まで、ロジットで推計。



Equation: UNTITLED Workfile: UNTITLED::Untitled*

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: BI
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 09/27/05 Time: 17:07
Sample: 1973M01 2002M01
Included observations: 349
Convergence achieved after 4 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.315278	0.153573	-2.052956	0.0401
IIP	0.214587	0.030497	7.036321	0.0000
KABU	0.031351	0.007355	4.262587	0.0000

Mean dependent var	0.578797	S.D. dependent var	0.494461
S.E. of regression	0.403107	Akaike info criterion	0.991180
Sum squared resid	56.22323	Schwarz criterion	1.024319
Log likelihood	-169.9610	Hannan-Quinn criter.	1.004372
Restr. log likelihood	-237.5564	Avg. log likelihood	-0.486994
LR statistic (2 df)	135.1909	McFadden R-squared	0.284545
Probability(LR stat)	0.000000		

Obs with Dep=0	147	Total obs	349
Obs with Dep=1	202		

Path = m:\eviews DB = none WF = untitled

質的従属変数の推計は、最小二乗法ではできないので、最尤法（ML）を使う。これにともない、係数の有意性の検定には、t-Statisticではなく、z-Statisticを用いるが、解釈の仕方は同じである。

最下段の Obs with Dep=0 は、被説明変数がゼロのサンプル数、Obs with Dep=1 は被説明変数が1の場合のサンプル数を示す。

(4) 当てはまりを示す2つの指標 - マクファーデンのR²と的中率

推計がうまくいったかどうかを調べるには、マクファーデンのR²（McFadden R-squared）と的中率を見るのが一般的だ。

マクファーデンの R^2

決定係数と同じようにゼロから 1 の間の数値を示し、1 に近いほど当てはまりが高いと考える。ただ、かなり当てはまりがよくても 0.4 か 0.5 程度にしかならない。

説明変数を定数項だけにした場合に比べて、どの程度尤度が上がるかという指標を 0 から 1 に基準化したものである。

的中率

的中率とは、説明変数から計算された推計値を見て、0.5 よりも大きい場合は 1、0.5 以下の場合はゼロとみなし、実際の被説明変数の値とあっているかどうかの割合を示す。推計した際の統計量の中には表示されないのので、別メニューを開く。

[View] [expectation prediction tables]

上の段のは個々の推計値に関する的中率を判断したもので、下の段は個々の推計値を合計したものを基準にしている。通常、上段の結果をみればよい。

< 行の説明 >

P(Dep=1)≤C	推計値が C (通常 0.5) 以下
P(Dep=1)>C	推計値が C より大きい
Total	合計
Correct	推計値が C 以下で実績値が 0 の場合 推計値が C より多く実績値が 1 の場合
% Correct	的中率
% Incorrect	非的中率
Total Gain*	常に 1 と推計した場合に比べてどのくらいの的中率が改善したか
Percent Gain**	常に 1 と推計した場合の非的中率に比べて何%改善したか。 改善割合 (Total Gain/Constant Probability の%incorrect)
E(# of Dep=0)	実績値がゼロであるサンプル数の期待値 (1 - 推計値の和)
E(# of Dep=1)	実績値が 1 であるサンプル数の期待値 (推計値の和)

< 列の説明 >

Estimated Equation	実績値の値 (推計値による分類)
Constant Probability	実績値の値 (常に 1 だと推計した場合の分類)

EViews

File Edit Object View Proc Quick Options Window Help

Equation: UNTITLED Workfile: UNTITLED::Untitled

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: BI
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 09/27/05 Time: 17:07
Sample: 1973M01 2002M01
Included observations: 349
Prediction Evaluation (success cutoff C = 0.5)

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	91	37	128	0	0	0
P(Dep=1)>C	56	165	221	147	202	349
Total	147	202	349	147	202	349
Correct	91	165	256	0	202	202
% Correct	61.90	81.68	73.35	0.00	100.00	57.88
% Incorrect	38.10	18.32	26.65	100.00	0.00	42.12
Total Gain*	61.90	-18.32	15.47			
Percent Ga...	61.90	NA	36.73			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	90.80	56.20	147.00	61.92	85.08	147.00
E(# of Dep=1)	56.20	145.80	202.00	85.08	116.92	202.00
Total	147.00	202.00	349.00	147.00	202.00	349.00
Correct	90.80	145.80	236.59	61.92	116.92	178.83
% Correct	61.77	72.18	67.79	42.12	57.88	51.24
% Incorrect	38.23	27.82	32.21	57.88	42.12	48.76
Total Gain*	19.65	14.30	16.55			
Percent Ga...	33.94	33.94	33.94			

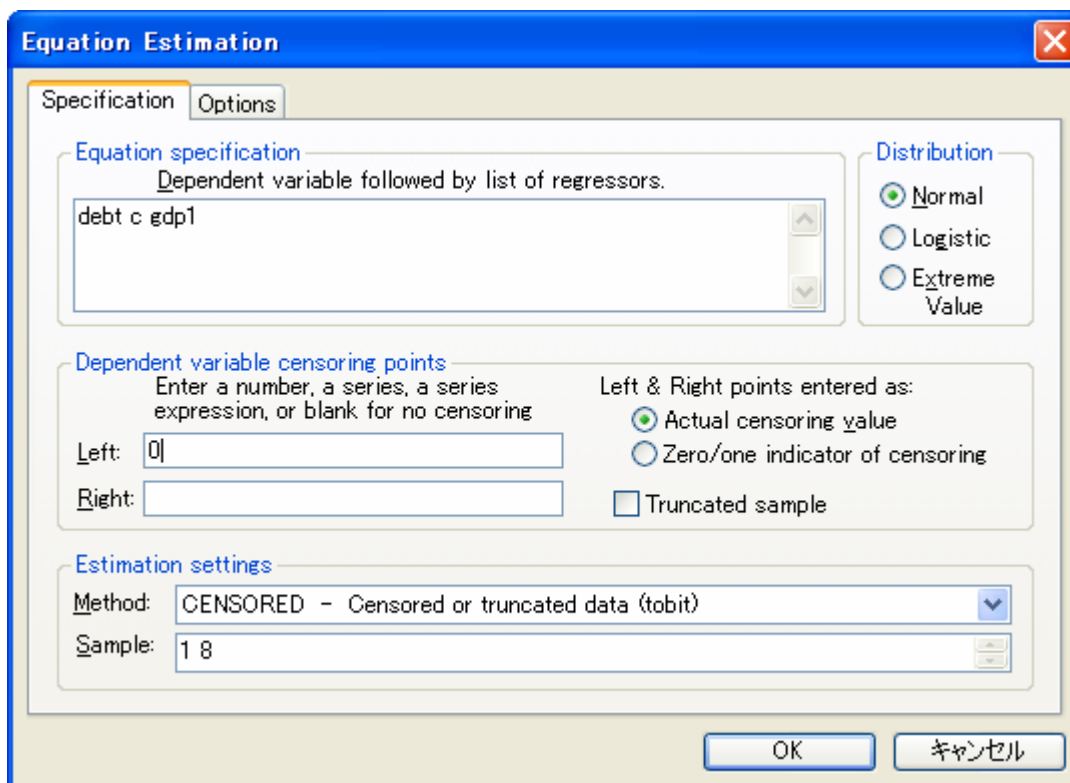
*Change in "% Correct" from default (constant probability) specification
**Percent of incorrect (default) prediction corrected by equation

Path = m:\eviews DB = none WF = untitled

最も重要なのは、上側の表の%CorrectのTotal(73.35%)で、推計値の的中率を表す。右側のConstant Probabilityは定数項に回帰したときの的中率であり、説明変数の値がいかなる値でも1であると予測したときの的中率を表している(57.88%)

すべてを1と予測した場合に比べて説明変数を使うことによってどれくらいの中率が上昇したかを表すのがTotal Gainで、この場合は15.47%ポイント(=73.35-51.24)上昇したことを示している。

トービット



Dependent Variable: DEBT
 Method: ML - Censored Normal (TOBIT) (Quadratic hill climbing)
 Date: 02/28/06 Time: 14:53
 Sample: 1 8
 Included observations: 8
 Left censoring (value) at zero
 Convergence achieved after 10 iterations
 Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
C	65.73642	4.160701	15.79936	0.0000
GDP1	-3.754633	0.633794	-5.924058	0.0000

Error Distribution				
SCALE:C(3)	6.550401	2.070500	3.163680	0.0016

R-squared	0.964736	Mean dependent var	30.15000
Adjusted R-squared	0.950630	S.D. dependent var	29.47318

S.E. of regression	6.548721	Akaike info criterion	4.873154
Sum squared resid	214.4287	Schwarz criterion	4.902944
Log likelihood	-16.49262	Hannan-Quinn criter.	4.672228
Avg. log likelihood	-2.061577		

Left censored obs	3	Right censored obs	0
Uncensored obs	5	Total obs	8
