一般化線形モデルを Excel で極め活用する

高橋行雄　（BioStat 研究所株式会社）

1. はじめに

一般化線形モデル（GLIM）の解を求めるためには，反復重み付き回帰，あるいは，情報行列を用いた反復計算が必要であり，現在，多くの統計ソフトに GLIM が実装されている．Excel には最小限の行列関数が備わっており，それらを用いて反復計算アルゴリズムを実装することができる．成書で GLIM の理論を学習する際に，自ら Excel で計算を行い，計算結果を再現することは，理論を正しく理解できているかを確認するために有益である．

古典的な生物検定法である 50 パーセント致死量（LD_{50}）を推定するためのプロビット法は，GLIM に包含されている．本発表では，GLIM を用いて LD_{50} を推定することを事例として取り上げる．歴史的にプロビット法は，計算時間の短縮のために反復重み付き回帰を適用する方法が多くの成書で示されてきた．田中ら訳（2008）には，ポアソン回帰の導入に反復重み付き回帰の事例が示されているが，プロビット法に関しては，情報行列を用いた結果のみが示されているだけであり，計算過程が明確に示されていない．また，他の成書でも，統計ソフトの使い方のマニュアル的になってしまっており，実用上の諸問題への応用方についての手順は示されていない．

プロビット法以外に低用量に裾を引く最小極値分布をリンク関数としてLD_{50}の推定を行う事例についても，Excel により計算を段階的に行なうことができる．さらに，環境ホルモンや毒性評価のためのベンチマークドーズとして世界的に使われている10パーセント致死量（LD_{10}）とその信頼区間についても，Excel による推定を行うことが可能である．このように様々な場面で Excel を用いた学習により，最尤法を使いこなすための応用力の向上が期待できる．

2. 反復重み付き回帰によるプロビット法

Finney（1971）により反復重み付き回帰を用いたプロビット法による LD_{50} の推定法が示され，佐久間（1977），吉村編著（1987）でも踏襲され，一般的な方法として普及してきた．この方法は，重み付き平方和を用いた計算方法で，重みおよび Probit 値を繰り返しごとに補正し，新に推定された推定値との差が近似的に 0 となった場合に収束したとみなす方法である．

GLIM では，期待値を $E(Y_i)=\mu_i$，リンク関数を $g(\mu_i)=\boldsymbol{x}^T\boldsymbol{\beta}=\eta_i$ としたときに，$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{z}$ による重み付き回帰を行い，Probit 値 z_i および重み w_{ii}

$$z_i=\eta_i+(y_i-\mu_i)\left(\frac{\partial\eta_i}{\partial\mu_i}\right),\qquad w_{ii}=\frac{1}{\mathrm{var}(Y_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2$$

を逐次変更し，収束するまで反復することでパラメータの推定を行う．出現率に対する用量反応関係にリンク関数として標準正規分布の逆累積分布 $\Phi^{-1}(\mu)$ に 5 を加えた Probit 値が伝統的に用いられている．吉村編著（1987）第 5.4 節のデータを用いて，伝統的なプロビット法の計算ではなく，GLIM で定式化されている行列計算により，反復重み付き回帰を Excel シート上に展開する．表 1 に 1 サイクル分の計算シートを示す．

表 1　プロビット法の初期値に対する重み付き回帰

	用量	切片	対数 dose	出現数	例数	出現率	probit 変換		回帰推定値	出現率推定値	補正 probit	重み	更新推定値	10）差
i	*dose*	x_0	x_1	r	n	y	z*		3) $\eta^{(1)}$	μ	4) z	5) w	9) $\eta^{(2)}$	$\eta^{(1)}$-$\eta^{(2)}$
1	101	1	2.004	0	10	0.000	-		3.522	0.070	3.001	2.762	3.253	0.270
2	136	1	2.134	2	10	0.200	4.158		4.241	0.224	4.161	5.146	4.081	0.160
3	183	1	2.262	5	10	0.500	5.000		4.957	0.483	5.000	6.362	4.907	0.050
4	247	1	2.393	8	10	0.800	5.842		5.681	0.752	5.833	5.367	5.743	-0.061
5	333	1	2.522	9	10	0.900	6.282		6.403	0.920	6.271	3.011	6.574	-0.172
6	450	1	2.653	10	10	1.000	-		7.130	0.983	7.531	1.045	7.413	-0.283
		$\boldsymbol{X}$				1) Intercept()=	-7.621	β_0^=	-7.621			8) 更新 β_0 ^ =	-9.598	2.1E-01
		デザイン行列				Slope()=	5.559	β_1^=	5.559			β_1^ =	6.411	平方和
						回帰係数	初期値	2) 係数・貼付				$(\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{X}^T\boldsymbol{Wz}$		
					6)	23.69	54.12	6)	7.363	-3.205	7)	119.58		
						54.12	124.34		-3.205	1.403		277.70		
						$\boldsymbol{X}^T\boldsymbol{WX}$			$(\boldsymbol{X}^T\boldsymbol{WX})^{-1}$			$\boldsymbol{X}^T\boldsymbol{Wz}$		

出現率 y が 0%と 100%の場合を除いた probit 値 $\boldsymbol{z}^*$ と対数用量 $\boldsymbol{x}_1$ から，1）Intercept および Slope 関数にて回帰係数の初期値を求め，それらを 2）［係数・貼付］欄に数値のみの貼付ける．3）推定値 $\boldsymbol{\eta}^{(1)}$ を Mmult 関数で $\eta_i^{(1)} = \boldsymbol{x}_i\boldsymbol{\beta}$ で計算し，出現率の推定値 μ を $\mu_i = \Phi(\eta_i - 5)$ を Normdis 関数で計算する．4）補正 probit 値を $z_i = \eta_i + (y_i - \mu_i)/\phi(\eta_i - 5)$ で計算す．5）重み w を $w_i = n_i\left[\phi(\eta_i - 5)\right]^2 / \left[\mu_i(1-\mu_i)\right]$ で計算する．ただし，$\phi(\eta_i - 5)$ は，標準正規分布の確率密度であり，Normdis 関数の関数形式：false で計算する．

重み付き回帰係数 $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{X}^T\boldsymbol{Wz}$ を計算するための重みは行列 $\boldsymbol{W}$ であるが，対角要素以外は 0 であり対角要素をベクトル $\boldsymbol{w}$ として Excel で行列とベクトルの計算 $\boldsymbol{X}*\boldsymbol{w}$ を計算し Transpose 関数で $(\boldsymbol{X}*\boldsymbol{w})^T$ を計算すると $\boldsymbol{X}^T\boldsymbol{W} = (\boldsymbol{X}*\boldsymbol{w})^T$ と同値になることを活用し，計算シートの簡潔化を図っている．6）さらに積和行列を $\boldsymbol{X}^T\boldsymbol{WX} = (\boldsymbol{X}*\boldsymbol{w})^T\boldsymbol{X}$ で計算し，Inverse 関数で分散共分散行列 $(\boldsymbol{X}^T\boldsymbol{WX})^{-1}$ を計算する．同様に 7）$\boldsymbol{X}^T\boldsymbol{Wz} = (\boldsymbol{X}*\boldsymbol{w})^T\boldsymbol{z}$ も別途計算し，8）重み付き回帰係数を行列の積を $\hat{\boldsymbol{\beta}}^{\text{new}} = \left[(\boldsymbol{X}^T\boldsymbol{WX})^{-1}\right]\left[\boldsymbol{X}^T\boldsymbol{Wz}\right]$ で計算し，その結果が［更新］した回帰係数として求められている．9）［更新］係数 $\hat{\boldsymbol{\beta}}^{\text{new}}$ を用いて推定値 $\boldsymbol{\eta}^{(2)}$ を $\eta_i^{(2)} = \boldsymbol{x}_i\boldsymbol{\beta}^{\text{new}}$ で計算し，さらに推定値 $\boldsymbol{\eta}^{(1)}$ との差 $\eta_i^{(1)} - \eta_i^{(2)}$ を計算した結果が示されている．ここまでが第 1 サイクルである．

10）差 $\eta_i^{(1)} - \eta_i^{(2)}$ が 0 とはみなせられず，平方和も大きいと判断されるので，$\hat{\boldsymbol{\beta}}^{\text{new}}$ をコピーして「係数・貼付」欄に値のみを張りつける．これが第 2 サイクル目で，$\boldsymbol{\beta}^{\text{new}} = [-10.041, 6.603]^T$ が得られる．同様の処理を行い第 3 サイクル目で $\boldsymbol{\beta}^{\text{new}} = [-10.069, 6.615]^T$ が得られる.同様に第 4 サイクル目とすると，差 $\eta_i^{(1)} - \eta_i^{(2)}$ が小数点以下 3 桁目はすべて 0 で，平方和は 5.3×10^{-8} となる．係数が小数点以下 3 桁で異なるので，第 5 サイクル目を行い，係数が小数点以下 3 桁まで一致し，平方和は 8.2×10^{-11} となるので繰返しを終了する．結果を表 2 に示す．

表 2 の回帰係数から 50%致死量が，$\log_{10} LD_{50} = (5-\hat{\beta}_0)/\hat{\beta}_1 = 2.278$，$LD_{50} = 10^{2.278} = 189.7$ として推定される．LD_{50} の近似信頼区間は，$(5-\hat{\beta}_0)/\hat{\beta}_1$ を $\hat{\beta}_0$ と $\hat{\beta}_1$ で偏微分し，これらを縦ベクトル $\boldsymbol{d}$ としたときに，分散が $\boldsymbol{d}^T(\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{d}$ となることから表 3 に示すように計算できる．

表 2　反復 5 回目の重み付き回帰によって得られた推定値

		1) Intercept()=	-7.621	β_0^=	-10.070			8) 更新 β_0^=	-10.070	8.2E-11
		Slope()=	5.559	β_1^=	6.615			β_1^ =	6.615	平方和
		回帰係数	初期値	2) 係数・貼付				$(\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{X}^T\boldsymbol{Wz}$		
	6)	20.64	47.13	6)	10.320	-4.498	7)	103.94		
		47.13	108.13		-4.498	1.970		240.72		
		$\boldsymbol{X}^T\boldsymbol{WX}$			$(\boldsymbol{X}^T\boldsymbol{WX})^{-1}$			$\boldsymbol{X}^T\boldsymbol{Wz}$		

表 3　LD_{50} および近似 95%信頼区間

	$(5-\beta_0$^$)/\beta_1$^	偏微分式	$\boldsymbol{d}$	$\boldsymbol{d}^T(\boldsymbol{X}^T\boldsymbol{WX})^{-1}\boldsymbol{d}$			対数	*dose*
対数LD_{50}=	2.278	d_0= -1/β_1^=	-0.151	*Var*=	0.0011	L95%=	2.213	163.3
LD_{50}=	189.7	-(5-β_0^)/β_1^2=	-0.344	*SE*=	0.0333	U95%=	2.343	220.5

3. 反復重み付き回帰によるロジット法および最小極値分布を用いる方法

出現率に対する用量反応曲線に正規分布の累積分布を用いるプロビット法に代わり，結果の解釈がしやすいロジスティック分布を用いたロジット法が使われるようになってきた．ロジット法を用いた LD_{50} とその信頼区間の推定する方法は，田中ら訳（2008）に示されている一般式から容易に導出できる．ロジット法では，出現率に 2 項分布を，リンク関数として $g(\mu_i)=\log[\mu_i/(1-\mu_i)]=\boldsymbol{x}^T\boldsymbol{\beta}=\eta_i$ としているので，補正値 z_i および重み w_{ii} を

$$z_i=\eta_i+(y_i-\mu_i)\left(\frac{\partial\eta_i}{\partial\mu_i}\right)=\eta_i+\frac{y_i-\mu_i}{\mu_i(1-\mu_i)},\qquad w_{ii}=\frac{1}{\mathrm{var}(Y_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2=\frac{n_i[\mu_i(1-\mu_i)]^2}{\mu_i(1-\mu_i)},$$

と導出することにより，LD_{50} とその信頼区間の推定がプロット法と同様に行なえる．

実際に表 1 の probit 変換 $\boldsymbol{z}^*$ を logit 変換 $z_i^*=\ln[y_i/(1-y_i)]+5$ に，出現率推定値 μ を $\mu_i=e^{\eta_i-5}/(1+e^{\eta_i-5})$ に，補正 probit z を補正 logit $z_i=\eta_i+(y_i-\mu_i)/[\mu_i(1-\mu_i)]$ に，重み w を $w_i=n_i\left[\mu_i(1-\mu_i)\right]^2/\left[\mu_i(1-\mu_i)\right]=n_i\left[\mu_i(1-\mu_i)\right]$ に変更する．初期値 $\hat{\boldsymbol{\beta}}=[-16.232, 9.357]^T$ からはじめて，第 4 サイクル目で $\boldsymbol{\beta}^{\mathrm{new}}=[-21.212, 11.523]^T$ に収束する．回帰係数は，プロビット法の場合と全く異なるが，$LD_{50}=188.2$，95%信頼区間は (161.5, 219.4) とほぼ同じ結果となる．

低用量の出現率が裾を引く場合には，最小極値分布を用量反応曲線とする必要がある．最小極値分布の累積分布は，$F_{SEV}(\mu)=1-\exp[-\exp(\beta_0+\beta_1 x)]$ であり，2 重に対数を取ることにより，リンク関数は，$g(\mu)=\ln[-\ln(1-\mu)]=\beta_0+\beta_1 x=\eta$ となり，補正値 z_i および重み w_{ii} を，

$$z_i=\eta_i+\frac{y_i-\mu_i}{-\ln(1-\mu_i)(1-\mu_i)},\qquad w_{ii}=\frac{n_i[-\ln(1-\mu_i)(1-\mu_i)]^2}{\mu_i(1-\mu_i)},$$

として，ロジット法と同様に計算できる．ロジット法の Excel シートを変更して，初期値 $\hat{\boldsymbol{\beta}}=[-9.214, 6.046]^T$ からはじめて，第 5 サイクル目で $\hat{\boldsymbol{\beta}}^{\mathrm{new}}=[-10.893, 6.743]^T$ に収束し，$LD_{50}=227.5$，95%信頼区間は（194.37, 266.16）と若干高めに推定される．

4. ロジット法に対する情報行列を用いた反復計算

反復重み付き回帰による最尤法の推定値は，情報行列を用いた最尤法と一致するが，分散共分散行列は微妙に異なる．反復重み付き回帰による最尤法は，GLIM に対してのみの適用であり，打ち切りデータがあるような場合には適用できない．情報行列を用いた最尤法は，このような制約がないが，複数のパラメータを持つモデル式の 1 階の偏微分ベクトル（スコア関

数）と，2 階の偏微分行列（ヘッセ行列 $\boldsymbol{H}$，情報行列 $\boldsymbol{I}$=−$\boldsymbol{H}$）を必要とする．

パラメータ数が 2 の場合は，Excel での行列計算が比較的容易に実施できる．パラメータ数が 3 以上の場合でも，パラメータの分散共分散行列（$\boldsymbol{I}^{-1}$）の算出を統計ソフトに委ねるならば，Excel に標準的に備わっているソルバーを用いて，対数尤度を直接最大化することで容易に最尤解を得ることができる．田中ら訳（2008）の「表 7.2 カブト虫の死亡データ」を用いて行列計算による反復計算を Excel シートで行なう．表 4 に初期値に $\boldsymbol{\beta}^{(m-1)}=[0, 0]^T$ を入力した結果を示す．

表 4　カブト虫の死亡数に対するロジット法による初期値での計算結果

	対数	死亡数	合計	出現率	推定値	対数尤度	行列 $\boldsymbol{U}$ の要素		情報行列 $\boldsymbol{I}$ の要素			
i	用量 x_i	y_i	n_i	p_i	π_i	$\ln(L_i)$	U_{i1}	U_{i2}	I_{i11}	I_{i12}	I_{i21}	I_{i22}
1	1.6907	6	59	0.102	0.500	-40.896	-23.500	-39.731	14.750	24.938	24.938	42.162
2	1.7242	13	60	0.217	0.500	-41.589	-17.000	-29.311	15.000	25.863	25.863	44.593
3	1.7552	18	62	0.290	0.500	-42.975	-13.000	-22.818	15.500	27.206	27.206	47.751
4	1.7842	28	56	0.500	0.500	-38.816	0.000	0.000	14.000	24.979	24.979	44.567
5	1.8113	52	63	0.825	0.500	-43.668	20.500	37.132	15.750	28.528	28.528	51.673
6	1.8369	53	59	0.898	0.500	-40.896	23.500	43.167	14.750	27.094	27.094	49.769
7	1.8610	61	62	0.984	0.500	-42.975	30.000	55.830	15.500	28.846	28.846	53.681
8	1.8839	60	60	1.000	0.500	-41.589	30.000	56.517	15.000	28.259	28.259	53.236
							50.500	100.785	120.25	215.71	215.71	387.43
切片 β_0^=		0.000		-37.856	-37.856	-333.40		50.500	120.25	215.71	6.743	-3.754
傾き β_1^=		0.000		21.337	21.337	$\ln L$		100.785	215.71	387.43	-3.754	2.093
		$\boldsymbol{\beta}^{(m-1)}+\left[\boldsymbol{I}^{(m-1)}\right]^{-1}\boldsymbol{U}^{(m-1)}=\boldsymbol{\beta}^{(m)}$						$\boldsymbol{U}^{(\mathrm{m-1})}$	$\boldsymbol{I}^{(\mathrm{m-1})}$		$\boldsymbol{I}^{-1}$分散共分散	

初期値に対し増分が $\left[\boldsymbol{I}^{(m-1)}\right]^{-1}\boldsymbol{U}^{(m-1)}=[-37.856, 21.337]^T$ であり，$\boldsymbol{\beta}^{(m-1)}$ にこの増分を加えて $\boldsymbol{\beta}^{(m)}$ とし，この結果を $\boldsymbol{\beta}^{(m-1)}$ に値のみを張りつける．このようにして反復計算を継続する．第 5 回目の $\boldsymbol{\beta}^{(m-1)}$ に対する $\boldsymbol{\beta}^{(m)}$ が同じ値となり収束した．なお，対数尤度= -186.24 であった．

表 5 反復計算によって得られたパラメータ

	初期値	1回目	2回目	3回目	4回目	5回目	6回目
切片 β_0^=	0.000	-37.856	-53.853	-59.965	-60.708	-60.717	-60.717
傾き β_1^=	0.000	21.337	30.384	33.844	34.265	34.270	34.270

5. ロジット法による LD_{10} と近似の 95%信頼区間の推定

情報行列を用いた反復計算で求められた結果から LD_{50} とその近似信頼区間を求める方法は，反復重み付き回帰の場合と同様に，LD_{50} の分散 $\boldsymbol{d}^T(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{d}$ に代えて $\boldsymbol{d}^T\boldsymbol{I}^{-1}\boldsymbol{d}$ として表 3 と同様に求めることができる．　反応が p_0 となるような対数用量 x_{p_0} は，

$$p_0=\frac{1}{1+\exp\left[-(\hat{\beta}_0+\hat{\beta}_1 x_{p_0})\right]}$$

で与えられる．この式を x_{p_0} について解き p_0 パーセント致死量となる用量は，

$$x_{p_0}=\frac{-\hat{\beta}_0+\ln(p_0/(1-p_0))}{\hat{\beta}_1}$$

となる．この分散は，パラメータによる偏微分を

$$d_0 = \frac{\partial x_{p_0}}{\partial \hat{\beta}_0} = -\frac{1}{\hat{\beta}_1}, \quad d_1 = \frac{\partial x_{p_0}}{\partial \hat{\beta}_1} = \frac{\hat{\beta}_0 - \ln(p_0/(1-p_0))}{\hat{\beta}_1^2}$$

として，これら用いた 2 次形式でより求められる．表 6 に第 5 回目の反復で得られた結果を元に，LD_{10} を求めると 1.708 となり，その 95%信頼区間は，（1.694，1.722）と計算される．

表 6　死亡率が 10 パーセントとなるロジット法による用量の推定

切片 β_0^=		-60.717	0.000	-60.717	-186.24		0.000	58.48	104.01	26.840	-15.082
傾き β_1^=		34.270	0.000	34.270	$\ln L$		0.000	104.01	185.09	-15.082	8.481
	$\boldsymbol{\beta}^{(m-1)} + \left[\boldsymbol{I}^{(m-1)}\right]^{-1} \boldsymbol{U}^{(m-1)} = \boldsymbol{\beta}^{(m)}$						$\boldsymbol{U}^{(\text{m-1})}$	$\boldsymbol{I}^{(\text{m-1})}$		$\boldsymbol{I}^{-1}$分散共分散	
		P=	0.10		$\boldsymbol{d}$		$\boldsymbol{d}^T(\boldsymbol{I})^{-1}\boldsymbol{d}$		x	*dose*	
		対数LD_p=	1.708	d_0=	-0.029	Var=	5.1E-05	L95%=	1.694	49.39	
		LD_p=	51.004	d_1=	-0.050	SE=	7.1E-03	U95%=	1.722	52.67	

6．LD_{10} の正確な 95%信頼区間

ロジスティック曲線の 95%信頼区間の算出方法を示す．回帰直線 $\hat{\eta}_i = \boldsymbol{x}_i \hat{\boldsymbol{\beta}}$ の分散は，$Var(\eta_i) = \boldsymbol{x}_i \boldsymbol{I}^{-1} \boldsymbol{x}_i^T$ であり，表 7 に示すように $\hat{\eta}$ とその 95%信頼区間を求める．その結果を $\pi_i = 1/[1+\exp(-\eta_i)]$ に適用して，出現率に対する 95%信頼区間を計算する．

表 7　ロジスティック曲線の 95%信頼区間の計算

		用量	死亡数	合計	出現率	推定値	回帰直線について				πについて	
i	x_0	x_1	y_i	n_i	p_i	π_i	η_i	分散	L95%	U95%	L95%	U95%
1	1	1.6907	6	59	0.102	0.059	-2.777	0.082	-3.339	-2.214	0.034	0.098
2	1	1.7242	13	60	0.217	0.164	-1.629	0.042	-2.030	-1.227	0.116	0.227
:												
8	1	1.8839	60	60	1.000	0.979	3.844	0.111	3.190	4.499	0.960	0.989

実際，$i=1$ の場合 $\boldsymbol{x}_1 = [1 \quad 1.6907]$ であり，$\hat{\eta}_1 = \boldsymbol{x}_1 \hat{\boldsymbol{\beta}} = \text{-}2777$ となり，$Var(\eta_1) = 0.082$，L95%= –2.777–1.96$\sqrt{0.082}$ = -3.339 から，π_1 については $\pi_{L95\%} = 1/[1+\exp(3.3339)] = 0.034$ となる．

LD_{10} について正確な 95%信頼区間は，η に関する回帰直線の 95%信頼区間を考慮して求めることができる．竹内（1979）によれば，95%信頼区間の下限の推定値を $\hat{x}_{L95}$ としたときに，$\hat{x}_{L95}$ における回帰直線上の推定値は，$\hat{\eta}_{\hat{x}_{L95}} = \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_{L95}$ で与えられる．その推定値 $\hat{\eta}_{\hat{x}_{L95}}$ の 95%信頼区間は，$\hat{x}_{L95}$ における 95%信頼区間 $= \hat{\eta}_{\hat{x}_{L95}} \pm t_\alpha \sqrt{Var(\hat{\eta}_{\hat{x}_{L95}})}$ で与えられる．その上側の 95%信頼曲線が，x_0 の逆推定値 η_0 を通る X 軸に平行な直線との交点は，$(\hat{\beta}_0 + \hat{\beta}_1 \hat{x}_{L95}) + t_\alpha \sqrt{Var(\hat{\beta}_0 + \hat{\beta}_1 \hat{x}_{L95})} = \eta_0$ で与えられる．この式を $\hat{x}_{L95}$ について解くことにより，$\hat{x}_0$ の下側の 95%信頼区間の式が求められる．

$$\left[Var(\hat{\beta}_0) - \frac{(\eta_0 - \hat{\beta}_0)^2}{t_\alpha^2}\right] + \left[2Cov(\hat{\beta}_0, \hat{\beta}_1) + \frac{2(\eta_0 - \hat{\beta}_0)\hat{\beta}_1}{t_\alpha^2}\right]\hat{x}_{L95} + \left[Var(\hat{\beta}_1) - \frac{\hat{\beta}_1^2}{t_\alpha^2}\right]\hat{x}_{L95}^2 = 0$$

式は複雑であるが，$\hat{x}_{L95}$ に関して 2 次式 $a + b\hat{x}_{L95} + c\hat{x}_{L95}^2 = 0$ となる．従って，2 次式の解の公式により x_{L95} を求めることができる．解は 2 つあるが，小さい方が $\hat{x}_{L95}$ となり，大きい方が $\hat{x}_{U95}$ となる．表 6 の $\boldsymbol{I}^{-1}$ より，$Var(\hat{\beta}_0) = 26.840$，$Var(\hat{\beta}_1) = 8.481$，$Cov(\hat{\beta}_0, \hat{\beta}_1) = -15.082$ を用い，$t_\alpha = 1.96$ として 95%信頼区間は，（1.691，1.720）と計算される．

p_0=	0.10	η_0=	-2.197	LD_{p0}=	1.708	L95%=	1.691	$\hat{x}_{L95} = \dfrac{-b \pm \sqrt{b^2-4ac}}{2c}$
a =	-864.6	b =	1013.9	c =	-297.2	U95%=	1.720	

7. 最小極値分布を用いた場合の LD_{10} と正確な 95%信頼区間の推定

反復重み付き回帰により最小極値分布を用いる場合については，すでに示した．情報行列を用いる場合には，次に示す偏微分式を用いる．

$$U_{i1} = \frac{\partial \ln L_i}{\partial \beta_0} = \frac{y_i(1-\pi_i)z_i}{\pi_i} - (n_i - y_i)z_i, \quad U_{i2} = \frac{\partial \ln L_i}{\partial \beta_1} = U_{i1}x_i, \text{ ただし } z_i = \exp\left(\beta_0 + \beta_1 x_i\right)$$

$$I_{i11} = -\frac{\partial^2 \ln L_i}{\partial \beta_0 \beta_0} = (n_i - y_i)z_i - \frac{y_i(1-\pi_i)(\pi_i - z_i)z_i}{\pi_i^2}, \quad I_{i12} = I_{i11}x_i, \quad I_{i21} = I_{i12}, \quad I_{i22} = I_{i11}x_i^2$$

これらの計算式を表 4 のロジット法での Excel シートを変更することにより，最小極値分布を用いた計算が行なえる．反復計算により対数尤度= -182.343， $\boldsymbol{\beta}^{(m)} = [-39.572, 22.041]^T$， $LD_{10} = 1.693$， 正確な 95%信頼区間（1.672,1.709）が得られる．

なお，表 2~7 で示した Excel での計算結果は，統計ソフト JMP の「モデルのあてはめ」でサポートされている「一般化線形モデル」を用いた結果と一致すること確認した．

8. 考察

LD_{50} のプロビット法による推定は，Finney（1971）により定式化され，現在に至っているものの，各種の平方和を用いた反復重み付き回帰の計算方法であり，LD_{50} の正確な 95%信頼区間の推定は，Filler の式による計算式が示されている．そのために，プロビット法をロジット法に変えた場合，あるいは，最小極値分布を用いた LD_{10} の推定などについて定式化するのが困難であった．

GLIM により，プロビット法，ロジット法あるいは最小極値分布について統一的な扱いが可能なので，Excel の行列関数を用いて GLIM による反復重み付き回帰および情報行列を用いた反復計算を 1 枚の Excel シート上に展開することにした．

同一シート上で推定された回帰パラメータと分散共分散行列を用いて LD_{50} および LD_{10} の近似 95%信頼区間および正確な 95%信頼区間の計算結果を示した．なお，この計算方法は，プロビット法でもロジット法でもすべて共通であり，汎用的に用いることができる．

Excel による情報行列を用いた最尤法は，GLIM の枠に留まらず非線形の問題，右側，左側および区間打ち切りがある場合の回帰分析などにも適用することが可能であり，統計ソフトが対応していない問題について容易に計算することができる．

文 献

Dobson A.J.著（2002）田中豊，森川敏彦，中山竹春ら訳（2008），一般化線形モデル入門，73-6，139-48，原著第 2 版，共立出版．

Dobson A.J.,Barnett A.G.(2008）,An Introduction to Genaralozed Linear Models 3rd ed.,CRC Press.

Finney D.J (1971), Probit Analysis 3rd ed.,50-80, Cambridge University Press.

佐久間昭（1977），薬効評価－計画と解析 I，330-9，東大出版会．

吉村功編著（1987），毒性･薬効データの統計解析，226-32，サイエンティスト社．

竹内啓（1979），数理統計学，312-5，東洋経済新報社．

連絡先：高橋行雄，〒105-0014 東京都港区芝 1-12-3 エステムプラザ芝公園 1005
BioStat 研究所株式会社，TEL/FAX 03-3452-8035，E-mail: takahashi.stat@nifty.com