第 2 期 医薬安全性研究会　基礎セミナー
◇ 日時：2012 年 6 月 1 日　（土）　10:00～12:00
◇ 場所：日本科学技術連盟、1 号館、3 階 講堂
協賛：サイエンティスト社

「じっくり勉強すれば身につく統計解析」を副題としたシリーズ全 3 巻がサイエンティスト社から刊行され，タイトルは「医薬品開発のための統計解析，第 1 部 基礎，第 2 部 実験計画法，第 3 部 非線形モデル」（芳賀グリーン本）である．本シリーズは，「医薬品開発のための統計解析講座」（SAS（株）JMP ジャパン事業部主催，年 12 回）のテキストとして使用されている. この本を題材として, 基礎セミナーを, 第 1 回目:「基本統計量とデータの比較」, 第 2 回目 :「回帰モデル」, 第 3 回目 :「共分散分析」, 第 4 回目 :「多重比較」, 第 5 回目 :「ロジスティック回帰」, 第 6 回目 :「効力比」と題して開催してきた.

第 7 回目は. 医薬品開発のための統計解析 第 1 部 4.4 節（5）逆推定, 第 3 部 1.2 節（8）非線形最小 2 乗法による標準誤差, で取り上げられている問題に対して, 回帰分析の基礎の基礎に戻ってじっくり解説する.

回帰分析の一歩進んだ活用として逆推定がある. 未知検体の測定値 y から濃度 x を求める場合, 回帰直線が求まれば y から x の点推定値は計算できるが, 推定精度を知るために必要な信頼区間の算出方法に触れた資料は少ない. Y 軸方向の平均値や個別データの 95%信頼区間は多くの統計ソフトで計算可能だが, 逆推定の個別データの 95%信頼区間までは対応していない. 本セッションでは, 使う機会は多いがキチンと解説されているとはいえない回帰分析について, Excel を用いて理解することを目的とする. そのために様々な応用が可能な計算の方法をじっくりと解説し, 数値例や手順の資料を提供する. そして, 受講者が回帰分析の逆推定や重み付き回帰分析などの高度の解析を日常的に使いこなせるようになることを願っている.

回帰分析・再入門

統計ソフトが対応していない生物統計の各種の課題をExcelでサクサク解こう

BioStat研究所（株）
高橋 行雄

1. はじめに

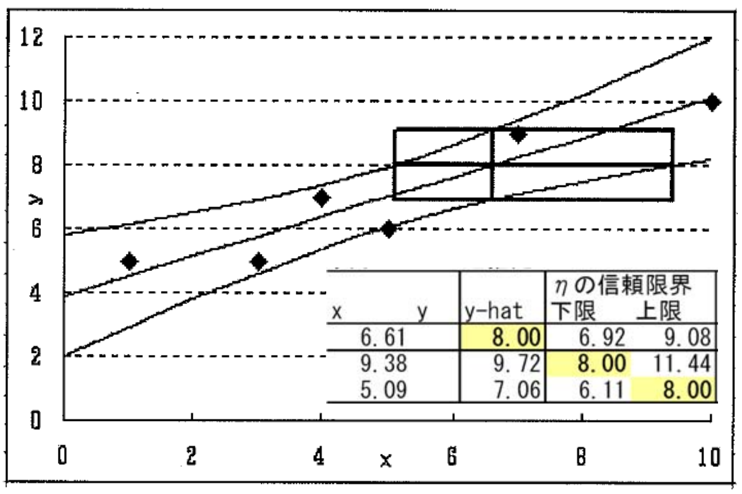
このスライドは次の電子図書の第2章の解説である
http://www.yukms.com/biostat/takahasi2/rec/003.htm
続・高橋セミナー 第3回 応用回帰分析1
－各種の重み付き回帰における逆推定－

用量反応関係での応用問題

- 回帰分析は，用量反応関係を想定した場合に多くの統計の教科書で取り上げられている分析方法で，（切片，傾き）の（推定値と *SE*）や（*t* 値，*p* 値）の計算まではどんな統計ソフトでも出力できる.
- しかし，そこから先の生物統計的な応用問題については，一般的な統計の教科書では解説されていない.

芳賀グリーン本 第3部，p22

表示 1.2.19　逆推定値 x_8 の信頼限界の意味



x	y	y-hat	η の信頼限界 下限	上限
6.61		8.00	6.92	9.08
9.38		9.72	8.00	11.44
5.09		7.06	6.11	8.00

- ある化合物Aの10 mg/kgの平均反応が *y*=8 であった.
- T薬の用量反応直線から，反応が *y*=8 となる用量とその95%信頼区間を求めたい.
- 個別データの95%信頼区間も求めたいのだが.

逆推定値の95%信頼区間 1

- Excel のソルバーを用いる方法
  - 回帰直線の95%信頼区間を求める式を用意
  - *x* を変化させて95%CLが *y*=8 となる x_8 をソルバーで探索する.

表示4.4.5 逆推定

x	y	y-hat	ηの信頼限界 下限	ηの信頼限界 上限	yの信頼限界 下限	yの信頼限界 上限
6.61		8.00	6.92	9.08	5.44	10.56
9.38		9.72	8.00	11.44	6.84	12.60
5.09		7.06	6.11	8.00	4.55	9.56
12.13		11.42	8.90	13.94	8.00	14.84
2.34		5.35	4.07	6.64	2.71	8.00

芳賀グリーン本 第1部 p206

逆推定値の95%信頼区間 2

- 個別データの95%信頼区間の式

$$(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95\,I})+t_\alpha\sqrt{Var(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95\,I})+\hat{\sigma}_e^2}=y_0$$

- この式を $\hat{x}_{L95}$ について解くと 2次式となるので, 2次式の解の公式で解けばよい.
- 生物検定法の常套手段であるが, どのような計算式になるのだろうか. また, 参考文献はどこにあるのだろうか.

重み付き回帰での逆推定

- 重み付き回帰分析で，逆推定値について個別データ95%信頼区間をどのように求めたらよいのだろうか.
- 文献は，見出すことができない.
- いつも頼りにしている JMP でも計算不可.
- Excel の行列関数を用いて回帰分析を行ない，各種の応用問題を解くことにする.

矩形データのままでの計算

- まず，Excelでデータを矩形（長方形，行列）に入力し．回帰分析のパラメータ（切片，傾き）を出すためのExcelで矩形データをそのまま扱う関数の使い方を丁寧に説明する.
- さらに，パラメータの分散と共分散を同時に計算する方法を丁寧に説明する.

共分散の活用

- パラメータ間の「共分散」を活用し回帰直線の95%信頼区間と個別データの95%信頼区間を計算する.
- 共分散を使うと統計の教科書にある煩雑な式を用いることなく回帰直線の95%信頼区間を簡単に計算できる.
- ただし，計算には矩形データのまま計算する作法（行列関数）の活用が必須です.

複数個のデータの平均の95%CL

- 個別データの95%信頼区間の計算は，教科書に出ているが，複数個のデータの平均の95%信頼区間は，どのようにして求めるのだろうか.
- 私もExcelでの計算作法の習得以前には，実際に計算したことはありませんでした.

2. 矩形データのままでの計算

始めの 第 1 歩，積和の計算

	A	B	C	D	E	F	G	H	I	J
1										
2		x								
3		x0	x1			β				
4		1	3			β0	-1			
5						β1	2			
6										
7		足し算，掛け算，割り算				積和の計算				
8		x0 + x1=	4	=B4+C4		x0*β0 + x1*β1=	5	=B4*G4 +C4*G5		
9										
10		x0 * x1=	3	=B4*C4		xβ =	5	=MMULT(B4:C4,G4:G5)		
11								（行方向，列方向）		

行方向のデータと列方向のデータの各セルの積を計算し，その和（合計）を計算する. MMULT（）関数を使う. 式を作成したら，最後にShift とCtrlキーを同時に押しながらEnterする.

推定値 *y*^の計算

	A	B	C	D	E	F	G	H	I
3									
4		*X*	*x*		β^		*y*^		
5		1	1		−1	=	1	=MMULT(B5:C9,E5:E6)	
6		1	1		2		1	:	
7		1	2				3	:	
8		1	3				5	:	
9		1	3				5	=MMULT(B5:C9,E5:E6)	
10									
11			(5×2)		(2×1)		(5×1)		

一般的に回帰分析では，切片に対して内部で変数“1”を付加している．これを明示的に与え，5行2列の矩形データとし，行ごとに(切片，傾き)の積和を計算し，*y* の推定値の計算を行なう．計算式はフィルハンドル でコピーする．

誤差平方和の計算

	A	B	C	D	E	F	G	H
2								
3								
4		*y*	−	*y*^	=	ε		
5		0		1		−1	=B5:B9−D5:D9	
6		2		1		1	=B5:B9−D5:D9	
7		3		3		0	=B5:B9−D5:D9	
8		4		5		−1	=B5:B9−D5:D9	
9		6		5		1	=B5:B9−D5:D9	
10								
11				2乗和	=	4	=SUMSQ(F5:F10)	
12								

y と *y*^の差を矩形データ全体を対象に計算する．F5 から F9 セルを選択したまま， *y* の範囲 B5:B9全体から*y*^の全体F5:F9を引く．ShiftとCtrlキーを同時に押しながらEnterする．誤差平方和は，SUMSQ()関数を用いて計算する．

矩形データの行と列の入れ替え

	A	B	C	D	E	F	G	H	I
2									
3		矩形データの行と列の入れ替え							
4		X			X^T				
5		1	1		1	1	1	1	1
6		1	1		1	1	2	3	3
7		1	2						
8		1	3		=TRANSPOSE(B3:C7)				
9		1	3			Shist + Ctrl + Enter			
10									
11		(5x2)			(2x5)				

(5×2)の矩形データ X の列と行を入れ替えて(2×5)の矩形データ X^T を TRANPOSE()関数を用いて作成してみよう.

矩形データ全体での積和

	A	B	C	D	E	F	G	H	I	J	K	L
2												
3											=MMULT(B5:F6,H5:I9)	
4		X^T						X		=	X^TX	
5		1	1	1	1	1		1	1		5	10
6		1	1	2	3	3		1	1		10	24
7								1	2		積和	
8								1	3		1行1列	1行2列
9								1	3		2行1列	2行2列
10												
11				(2x5)				(5x2)			(2x2)	

(2×5)の矩形データ X^T と(5×2)の矩形データ X の積和を考えよう. 最初のX^T の全ての行(2行)と 次のX の全ての列(2列)の組み合わせに対して積和をMMULT()関数で同時に計算する.

逆数

	A	B	C	D	E	F	G	H	I	J
1										
2							(2×2)の矩形データ			
3			n				A			
4			5				5	0	列方向	
5							0	20		
6		逆数 1/n =	0.2	=1/C4			A^{-1}			
7						“逆数”	0.2	0	行方向	
8			$n\times(1/n)=1$				0	0.05		
9										
10						$A^{-1}A$ =	1	0	=MMULT(G7:H8,G4:H5)	
11							0	1		

***n*=5 の逆数*n*$^{-1}$は *n*$^{-1}$= 1/*n* =1/5.（2×2）の矩形データ *A* の逆数 *A*$^{-1}$は？　*A* と *A*$^{-1}$の 積和*A A*$^{-1}$の対角要素が全て 1 となるような矩形データと定義されている.**

矩形データの逆数

	A	B	C	D	E	F	G	H	I	J	K
3											
4		X^TX			$(X^TX)^{-1}$		=	$(X^TX)\ (X^TX)^{-1}$			
5		5	10		?	?		1	0	定義より	
6		10	24		?	?		0	1	対角要素が全て1	
7											
8					=MINVERSE(B5:C6)			=MMULT(B5:C6,E9:F10)			
9					1.2	−0.5		1	0		
10					−0.5	0.25		0	1		
11											

元の矩形データ（*X*^{T}*X*）に掛けて，対角要素が全て 1 となる矩形データは，MINVERSE（）関数で計算できる．MMULT（）関数で検算をすると対角要素が全て1となる.

回帰パラメータの計算

- ◆矩形データ X を転置： X^T
- ◆X^T と X の積和： $(X^T X)$
- ◆積和 $(X^T X)$ の**逆行列**： $(X^T X)^{-1}$
- ◆更にX^T との積和： $(X^T X)^{-1} X^T$
- ◆反応 y との矩形データとの積和：
  $(X^T X)^{-1} X^T y$
- ◆パラメータの推定： $\hat{\beta} = (X^T X)^{-1} X^T y$

Excel による段階的な計算

積和はMMULT()関数

X^T (2x5)

1	1	1	1	1
1	1	2	3	3

X (5x2)

1	1
1	1
1	2
1	3
1	3

= $X^T X$ (2x2)

5	10
10	24

$(X^T X)^{-1}$ (2x2)

1.2	−0.5
−0.5	0.25

X^T (2x5)

1	1	1	1	1
1	1	2	3	3

= $(X^T X)^{-1} X^T$ (2x5)

0.7	0.7	0.2	−0.3	−0.3
−0.25	−0.25	0	0.25	0.25

$(X^T X)^{-1} X^T$ (2x5)

0.7	0.7	0.2	−0.3	−0.3
−0.25	−0.25	0	0.25	0.25

y (5x1)

0
2
3
4
6

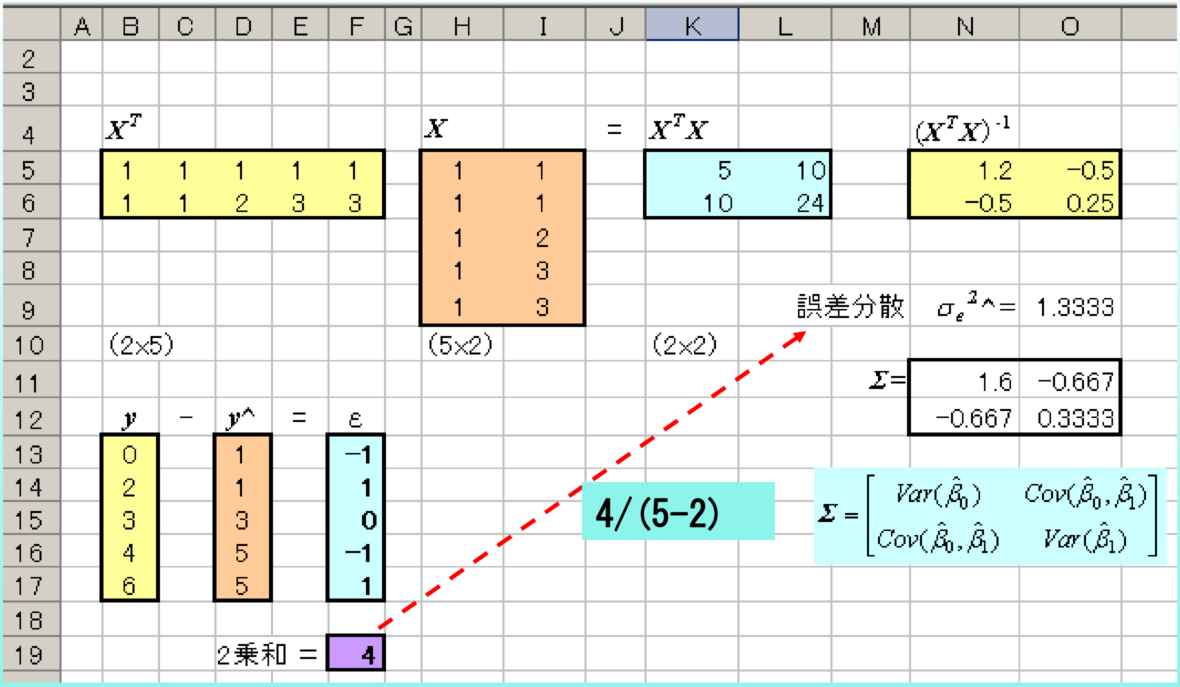
= β (2x1)

−1
2

切片と傾きの分散，共分散

- 推定値 y^を計算する．
- y と y^の差の平方和を S_T とする．
- データ数 n，推定したパラメータの数
- 誤差分散 $\sigma_e^2 = S_T / (n - p)$ で求める．
- 積和 (X^TX) の逆行列：$(X^TX)^{-1}$ に誤差分散 σ_e^2 を掛けて，分散共分散行列 Σ を計算．
- 対角要素が，切片 β_0 と傾き β_1 の分散となる．

Excel による分散共分散の計算



回帰パラメータの95%信頼区間

回帰パラメータ

k	$\boldsymbol{x}$	β_k	β_k^	$Var(\beta_k$^$)$	SE	$L95\%$	$U95\%$
0	x_0:切片	β_0	−1.0	1.6000	1.2649	−5.0255	3.0255
1	x_1:傾き	β_1	2.0	0.3333	0.5774	0.1626	3.8374

誤差分散 $\sigma_e{}^2$^ = 1.3333

$$\Sigma = \boxed{\begin{matrix} 1.6 & -0.667 \\ -0.667 & 0.3333 \end{matrix}}$$

$$\Sigma = \begin{bmatrix} Var(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) \\ Cov(\hat{\beta}_0, \hat{\beta}_1) & Var(\hat{\beta}_1) \end{bmatrix}$$

分散共分散行列 Σ（矩形データ）の対角要素が，切片 β_0 と傾き β_1 の分散となる

2. 個別データの95%信頼区間

推定値の分散

	A	B	C	D	E	F	G	H	I	J	K
2											
3		X									
4	i	x_0	x_1		$Var(y\hat{})$						
5	1	1	1		0.6000	=MMULT(MMULT(B5:C5,$B$13:$C$14)					
6	2	1	1		0.6000	,TRANSPOSE(B5:C5))					
7	3	1	2		0.2667						
8	4	1	3		0.6000						
9	5	1	3		0.6000	=MMULT(MMULT(B9:C9,$B$13:$C$14)					
10						,TRANSPOSE(B9:C9))					
11			(5×2)								
12											
13	Σ=	1.6	−0.667								
14		−0.667	0.3333								
15											

$$Var(\hat{y}_i) = \mathbf{x}_i \boldsymbol{\Sigma} \mathbf{x}_i^T$$

え！　これは何ですか？

i = 5 の場合の推定値の分散

- i = 5

i	$\mathbf{x}_i$		Σ		$\mathbf{x}_i^T$	=	$Var(y_i\hat{})$
5	1	3	1.6	−0.667	1		0.6
			−0.667	0.3333	3		

- **分散に関する基本演算と同じ結果**

$$
\begin{aligned}
Var(\hat{y}_5) &= Var(\hat{\beta}_0 x_{0,5} + \hat{\beta}_1 x_{1,5}) \\
&= Var(\hat{\beta}_0 + 3\hat{\beta}_1) \\
&= Var(\hat{\beta}_0) + 2\times 3 Cov(\hat{\beta}_0, \hat{\beta}_1) + 3^2 Var(\hat{\beta}_1) \\
&= 1.600 - 6\times 0.667 + 9\times 0.333 = 0.600
\end{aligned}
$$

じっくり計算してみましょう

i	$\boldsymbol{x}_i$		$\boldsymbol{\Sigma}$		$\boldsymbol{x}_i^T$	=	$Var(y_i\hat{})$
5	1	3	1.6	-0.667	1		0.6
			-0.667	0.3333	3		

$\boldsymbol{x}_i\,\boldsymbol{\Sigma}$

		$\boldsymbol{x}_i^T$
1×1.6 + 3×(-0.667)	1×(-0.667) + 3×0.3333	1
		3

$\boldsymbol{x}_i\,\boldsymbol{\Sigma}\,\boldsymbol{x}_i^T$

= 1×{1×1.6 + 3×(-0.667)}+ 3×{1×(-0.667) + 3×0.3333}

= 1.6 + 3×(-0.667)}+ 3×(-0.667) + 3× 3×0.3333}

= 1.6 + 2×3×(-0.667)} + 3×3×0.3333}

矩形データの積は，（行方向･列方向）が基本

分散に関する基本演算

◆ **基本形**

$$\begin{aligned}
Var(\hat{y}_i) &= Var(\hat{\beta}_0 x_{0i} + \hat{\beta}_1 x_{1i}) \\
&= Var(\hat{\beta}_0 x_{0i}) + 2Cov(\hat{\beta}_0 x_{0i}, \hat{\beta}_1 x_{1i}) + Var(\hat{\beta}_1 x_{1i}) \\
&= x_{0i}^2 Var(\hat{\beta}_0) + 2x_{0i}x_{1i}Cov(\hat{\beta}_0, \hat{\beta}_1) + x_{1i}^2 Var(\hat{\beta}_1) \\
&= 1.600 - 6 \times 0.667 + 9 \times 0.333 = 0.600
\end{aligned}$$

◆ **一般式**

$$Var(\hat{y}_i) = \sum_{k=0}^{p-1}\sum_{k'=0}^{p-1} x_{ki}x_{k'i}Cov(\hat{\beta}_k, \hat{\beta}_{k'})$$

$$Cov(\hat{\beta}_0, \hat{\beta}_0) = Var(\hat{\beta}_0) \qquad Cov(\hat{\beta}_1, \hat{\beta}_1) = Var(\hat{\beta}_1)$$

◆ **矩形データ**

$$Var(\hat{y}_i) = \boldsymbol{x}_i \boldsymbol{\Sigma} \boldsymbol{x}_i^T$$

推定値の95%信頼区間

◆回帰直線の95%信頼区間

$$\hat{y}_i の95\%\mathrm{CL} = \hat{y}_i \pm t(f_e, 0.05)\sqrt{Var(\hat{y}_i)}$$

◆個別データの95%信頼区間

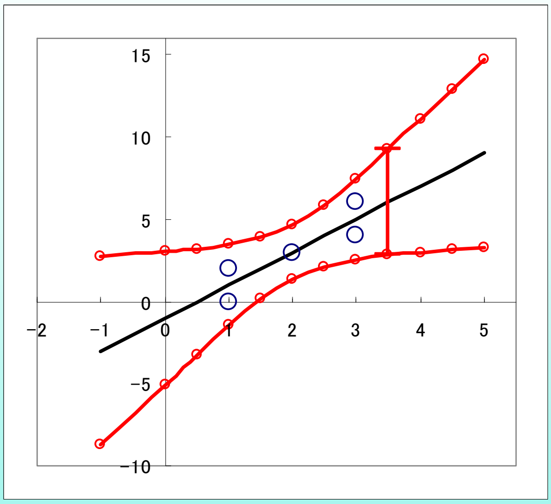
$$Var(\hat{y}_i + e_i) = Var(\hat{y}_i) + Var(e_i) = Var(\hat{y}_i) + \hat{\sigma}_e^2$$

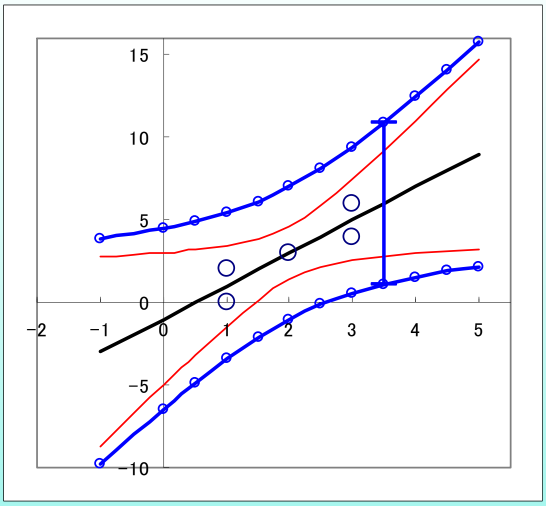
$$個別データの95\%CL = \hat{y}_i \pm t(f_e, 0.05)\sqrt{Var(\hat{y}_i) + \hat{\sigma}_e^2}$$

95%信頼区間の計算結果

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		回帰分析の基礎･応用					2012/09/15 高橋行雄										
2																	
3						$(X^TX)^{-1}$						$t_{0.05}$=	3.182	f_e=	3		
4		n=	5			1.2	-0.5					Σ=	1.600	-0.667			
5		p=	2			-0.5	0.25						-0.667	0.333			
6												σ_e^2^=	1.333				
7																	
8		i		y		X			β^		y^		$Var(y_i$^$)$	$L95R$	$U95R$	$L95I$	$U95I$
9		1		0		1	1		-1	=	1		0.600	-1.465	3.465	-3.425	5.425
10		2		2		1	1		2		1		0.600	-1.465	3.465	-3.425	5.425
11		3		3		1	2				3		0.267	1.357	4.643	-1.026	7.026
12		4		4		1	3				5		0.600	2.535	7.465	0.575	9.425
13		5		6		1	3				5		0.600	2.535	7.465	0.575	9.425

95%信頼区間の表示





回帰の95%CL　　　　個別データの95%CL

作図法の詳細は，電子図書の

2.6節「Excel による計算方法および作図方法」を参照のこと

作図のための計算表

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
7																	
8		*i*		*y*		*X*			β^		*y*^		Var(y_i^)	L95R	U95R	L95I	U95I
9		1		0		1	1		−1	=	1		**0.600**	**−1.465**	3.465	**−3.425**	5.425
10		2		2		1	1		2		1		0.600	−1.465	3.465	−3.425	5.425
11		3		3		1	2				3		0.267	1.357	4.643	−1.026	7.026
12		4		4		1	3				5		0.600	2.535	7.465	0.575	9.425
13		5		6		1	3				5		0.600	2.535	7.465	0.575	9.425
14																	
15		図作成のための				1	−2.0				−5.0		5.600	−12.531	2.531	−13.380	3.380
16		計算シート				1	−1.0				−3.0		3.267	−8.752	2.752	−9.826	3.826
17						1	0.0				−1.0		1.600	−5.026	3.026	−6.451	4.451
18						1	0.5				0.0		1.017	−3.209	3.209	−4.879	4.879
19						1	1.0				1.0		0.600	−1.465	3.465	−3.425	5.425
20						1	1.5				2.0		0.350	0.117	3.883	−2.129	6.129
21						1	2.0				3.0		0.267	1.357	4.643	−1.026	7.026
22						1	2.5				4.0		0.350	2.117	5.883	−0.129	8.129
23						1	3.0				5.0		0.600	2.535	7.465	0.575	9.425
24						1	3.5				6.0		1.017	2.791	9.209	1.121	10.879
25						1	4.0				7.0		1.600	2.974	11.026	1.549	12.451
26						1	4.5				8.0		2.350	3.121	12.879	1.892	14.108
27						1	5.0				9.0		3.267	3.248	14.752	2.174	15.826
28						1	6.0				11.0		5.600	3.469	18.531	2.620	19.380
29						1	7.0				13.0		8.600	3.667	22.333	2.970	23.030
30						1	8.0				15.0		12.267	3.854	26.146	3.264	26.736
31						1	9.0				17.0		16.600	4.034	29.966	3.523	30.477
32						1	10.0				19.0		21.600	4.209	33.791	3.760	34.240

3. 逆推定値の95%信頼区間

逆推定の点推定

◆元の回帰式

$$y_0 = \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_0$$

◆ $\hat{x}_0$ について解き，$\hat{\beta}_0 = -1,\ \hat{\beta}_1 = 2$　の場合

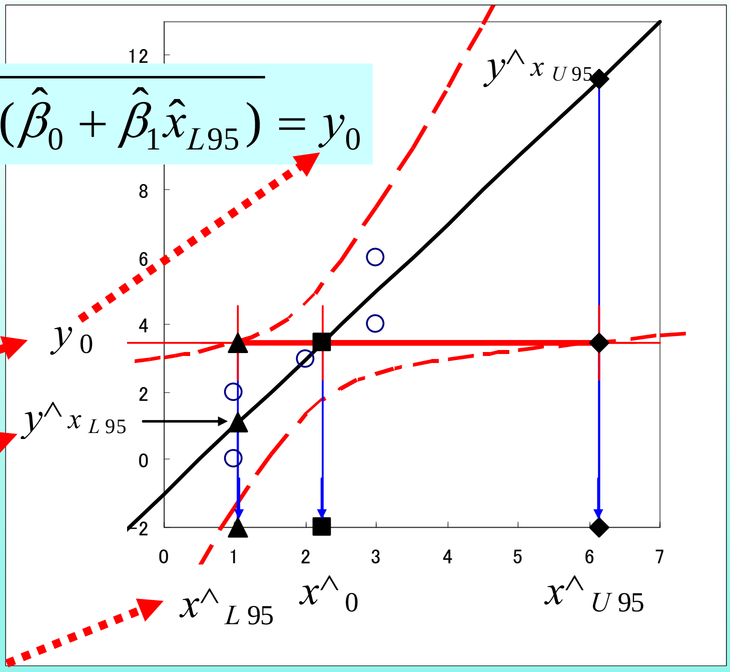
$$\hat{x}_0 = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1} = \frac{3.5-(-1)}{2} = 2.25$$

逆推定の95%信頼区間の方程式

$$(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95})+t_\alpha\sqrt{Var(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95})}=y_0$$

$$\hat{y}_{\hat{x}_{L95}}+t_\alpha\sqrt{Var(\hat{y}_{\hat{x}_{L95}})}$$

$$\hat{y}_{\hat{x}_{L95}}=\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95}$$



下限の推定値

$x\hat{}_{L95}$ について解く

$$Var(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95})=\left(\frac{y_0-\hat{\beta}_0-\hat{\beta}_1\hat{x}_{L95})}{t_\alpha}\right)^2$$

$$Var(\hat{\beta}_0)+2\hat{x}_{L95}Cov(\hat{\beta}_0,\hat{\beta}_1)+\hat{x}_{L95}^2Var(\hat{\beta}_1)=\frac{(y_0-\hat{\beta}_0)^2-2(y_0-\hat{\beta}_0)\hat{\beta}_1\hat{x}_{L95}+\hat{\beta}_1{}^2\hat{x}_{L95}^2}{t_\alpha^2}$$

$$\left\{Var(\hat{\beta}_0)-\frac{(y_0-\hat{\beta}_0)^2}{t_\alpha^2}\right\}+\left\{2Cov(\hat{\beta}_0,\hat{\beta}_1)+\frac{2(y_0-\hat{\beta}_0)\hat{\beta}_1}{t_\alpha^2}\right\}\hat{x}_{L95}+\left\{Var(\hat{\beta}_1)-\frac{\hat{\beta}_1{}^2}{t_\alpha^2}\right\}\hat{x}_{L95}^2=0$$

式は複雑であるが, $\hat{x}_{L95}$ に関して2次式

$$a+b\,\hat{x}_{L95}+c\,\hat{x}_{L95}^2=0$$

2 次式の解に帰着

$$a + b\hat{x}_{L95} + c\hat{x}_{L95}^2 = 0$$

$$a = Var(\hat{\beta}_0) - \frac{(y_0 - \hat{\beta}_0)^2}{t_\alpha^2}$$

$$b = 2Cov(\hat{\beta}_0, \hat{\beta}_1) + \frac{2(y_0 - \hat{\beta}_0)\hat{\beta}_1}{t_\alpha^2}$$

$$c = Var(\hat{\beta}_1) - \frac{\hat{\beta}_1^{\,2}}{t_\alpha^2}$$

$$\hat{x}_{L95} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2c}$$

Excel による計算

β＾

−1
2

$t_{0.05}$= 3.182　　f_e= 3　　$y_0 = 3.5$

Σ=	1.600	−0.667
	−0.667	0.333

$\sigma_e{}^2$ ^ = 1.333

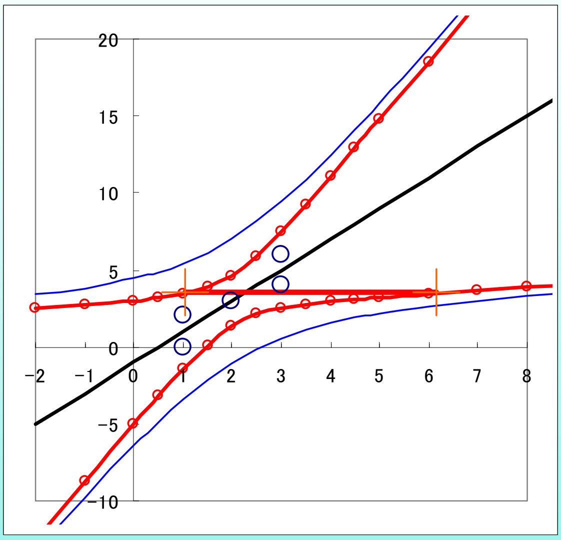
$$a = Var(\hat{\beta}_0) - \frac{(y_0 - \hat{\beta}_0)^2}{t_\alpha^2} = 1.6 - \frac{(3.5-(-1))^2}{3.182^2} = -0.399$$

$$b = 2Cov(\hat{\beta}_0, \hat{\beta}_1) + \frac{2(y_0 - \hat{\beta}_0)\hat{\beta}_1}{t_\alpha^2} = 2\times(-0.667) + \frac{2\times(3.5-(-1))\times 2}{3.182^2} = 0.444$$

$$c = Var(\hat{\beta}_1) - \frac{\hat{\beta}_1^{\,2}}{t_\alpha^2} = 0.333 - \frac{2^2}{3.182^2} = -0.062$$

$$(\hat{x}_{L95}, \hat{x}_{U95}) = \frac{-0.444 \pm \sqrt{0.444^2 - 4\times(-0.399)\times(-0.062)}}{2\times(-0.062)} = (1.054, 6.151)$$

回帰の95%CLに対する逆推定



逆推定 y_0 = 3.500　　x_0^ = 2.250

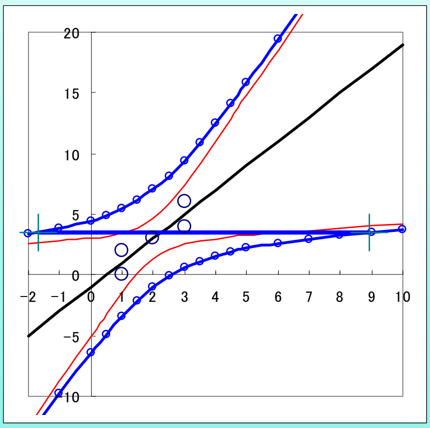
回帰直線の95%CL

a =	−0.399	$x_{L\,95R}$^ =	1.054
b =	0.444	$x_{U\,95}$^ =	6.151
c =	−0.062		

個別データの95%CLでは

$$(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95I})+t_\alpha\sqrt{Var(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95I})+\hat{\sigma}_e^2}=y_0$$

2次式の定数項 a に $\hat{\sigma}_e^2$ が加わる



個別データの95%CL

a' =	0.934	$x_{L\,95I}$^ =	−1.702
b =	0.444	$x_{U\,95I}$^ =	8.907
c =	−0.062		

m 個の平均の95%CL

$$\overline{y}_{i\cdot}=\frac{y_{i1}+\cdots+y_{im}}{m}=\frac{\hat{y}_{i1}+e_{i1}}{m}+\cdots+\frac{\hat{y}_{im}+e_{im}}{m}=\hat{y}_i+\frac{e_{i1}}{m}+\cdots+\frac{e_{im}}{m}$$

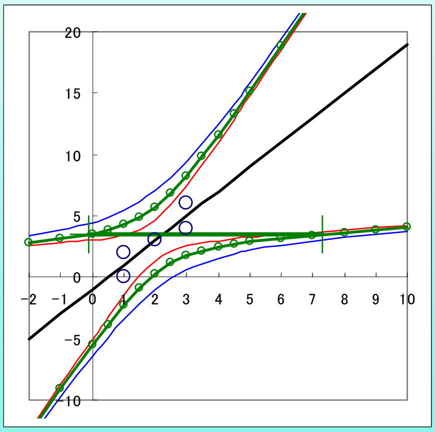
$$Var(\overline{y}_{i\cdot})=Var(\hat{y}_i+\frac{e_{i1}}{m}+\cdots+\frac{e_{im}}{m})=Var(\hat{y}_i)+\frac{1}{m^2}\hat{\sigma}_e^2+\cdots+\frac{1}{m^2}\hat{\sigma}_e^2=Var(\hat{y}_i)+\frac{\hat{\sigma}_e^2}{m}$$

$$m\text{個の平均の}95\%\,CL=\hat{y}_i\pm t(f_e,0.05)\sqrt{Var(\hat{y}_i)+\frac{\hat{\sigma}_e^2}{m}}$$

m 個の平均の逆推定

$$(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95I})+t_\alpha\sqrt{Var(\hat{\beta}_0+\hat{\beta}_1\hat{x}_{L95I})+\frac{\hat{\sigma}_e^2}{m}}=y_0$$

2次式の定数項 *a* に $\frac{\sigma_e^2}{m}$ が加わる

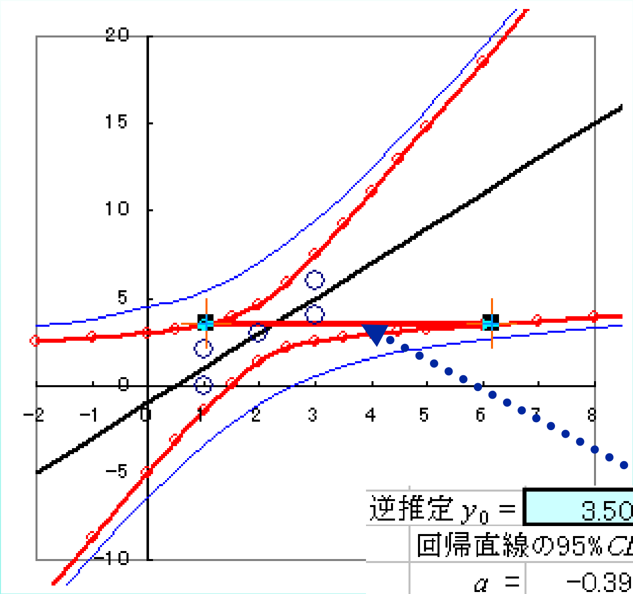


m 個の平均の95%CL

a'' =	0.045	x_{L95m}^ =	−0.100
b =	0.444	x_{U95m}^ =	7.305
c =	−0.062		

m=3 の場合

逆推定値の95%CLの作図



逆推定 y_0 =	3.500	x_0^ =	2.250	
回帰直線の95%CL				y
a =	−0.399	x_{L95R}^ =	1.054	3.50
b =	0.444	x_{U95}^ =	6.151	3.50
c =	−0.062			

- 【元のデータ】で【系列】を追加
- 【データ系列の書式設定／パターン／線】で線とマーカを設定

まとめ

最初の 1 歩

- 副題は「統計ソフトが対応していない生物統計の各種の課題をExcelでサクサク解こう」であった.
- 逆推定は,「生物統計」の課題の代表例である.
- SASでは, PROBITプロシジャで, 2 値反応でのシグモイド曲線の問題について逆推定はできるが, REGプロシジャなどの通常の回帰分析では, 対応していない.

JMP では逆推定ができるが

- JMPでは, 単回帰に関して, 逆推定ができる.
  - ただし, *m*個の平均に関しては対応していない.
- JMPでできるのならば, Excel を使う必要はないのではないか, と思っていた.
- しかし, 検量線などで一般的に使われている, 重み付き回帰では, 個別データの95%CLの計算で, 重みが考慮されていないことに, 気がついた.
- この問題解決のために, Excel を徹底的に使うことしかないと覚悟した.

単回帰から矩形データのまま

- 通常の回帰分析に関する書物では，Σを用いた計算式が示されていて，これは，手計算を前提にした時代の産物である.
- Eexcl でもΣを用いた計算は可能であるが，矩形データのままでの関数計算のほうが，生産的であり，さらなる応用のためには，最初から，矩形データのままでの計算を徹底的に追及した.

生物統計の各種の課題

- 単回帰の逆推定は，一つの応用
  - 計算過程の透明化が，今後の発展に不可欠
- さらなる課題（本日は割愛）
  - 勾配比による効力比
  - 平行線のあてはめによる効力比
  - 非線形回帰の教育用としてExcelを活用
- 目指すところ
  - 電卓を前提にした回帰分析に関する入門教育をExcel を前提にした新たな入門教育の提案