

第 9 回 続高橋セミナー
最尤法によるポアソン回帰分析入門
2020 年 4 月 27 日

第 2 章 ニュートン・ラフソン法によるポアソン回帰

「最尤法は，どのような方法なのか」との質問に「尤度を最大にする方法である」との紋切り型な説明がはびこっている．最小 2 乗法であれば，「誤差平方和を最小にする方法である」との説明と同様であり，まともな回答ではない．尤度とは何か，最大化するための方法はあるようなものか，Excel の行列関数を用いて Excel シート 1 枚の中で実現することによる可視化を試みる．まず，最尤法によるポアソン回帰のパラメータ推定について，第 1 章で取り上げた事例を用い，最尤法の基本となる対数尤度の計算を Excel の関数を補助的に用い，手作業で最大化するための手順を示す．次に，対数尤度関数をパラメータに関して偏微分した偏微分ベクトル，更に偏微分をした 2 階の偏微分行列（ヘッセ行列）の導出を丁寧に示す．それらを Excel シート上に展開し，ニュートン・ラフソン法による対数尤度の最大化の方法について示す．対数尤度の最大化の方法について，1 群の場合，回帰の場合，対数リンクでの回帰の場合，オフセットがある場合について第 1 章で取り上げた事例を用いて Excel を用いることにより計算過程を可視化しつつ示す．

第 2 章 目 次

2.	ニュートン・ラフソン法によるポアソン回帰	63
2.1.	手作業による逐次的な対数尤度の最大化	63
2.2.	Excel のソルバーによる対数尤度の最大化	68
2.3.	ニュートン・ラフソン法による対数尤度の最大化	70
2.4.	ポアソン回帰のバリエーション	78
2.5.	対数リンクの場合のポアソン回帰	84
2.6.	対数リンクでオフセットがある場合のポアソン回帰	88
	文献索引，索引，解析用ファイル一覧	95

第9回 続高橋セミナー 最尤法によるポアソン回帰分析入門

第9回 続高橋セミナー「最尤法によるポアソン回帰分析入門」は、ページ数が多いので章ごとに公開する。全体の章立てを次に示す。

目次

はじめに -----	1
1. ポアソン分布に従う各種のカウント・データ-----	7
2. ニュートン・ラフソン法によるポアソン回帰-----	63
3 尤度比検定のためのデザイン行列-----	95
4. デザイン行列を用いた回帰分析入門-----	135
5. 反復重み付き最尤法によるポアソン回帰 -----	175
6. 過分散・ゼロ過剰への対応 -----	207
7. 過分散がある場合の探索的ポアソン回帰-----	237
8. 2本の回帰直線の比較-----	269
9. 花数を共変量とした種子数の分析 -----	293
10. オフセットを含む探索的ポアソン回帰-----	321
11. デビアンس・逸脱度・残差・テコ比 -----	357
12. パラメータの共分散行列の活用 -----	379
13. 最小2乗平均の謎を予測プロファイルで解く -----	417
文献, 文献索引, 索引, 解析用ファイル一覧 -----	451

2. ニュートン・ラフソン法によるポアソン回帰

「最尤法は，どのような方法なのか」との質問に「尤度を最大にする方法である」との紋切り型な説明がはびこっている．最小 2 乗法であれば，「誤差平方和を最小にする方法である」との説明と同様であり，まともな回答ではない．尤度とは何か，最大化するための方法はあるようなものか，Excel の行列関数を用いて Excel シート 1 枚の中で実現することによる可視化を試みる．まず，最尤法によるポアソン回帰のパラメータ推定について，第 1 章で取り上げた事例を用い，最尤法の基本となる対数尤度の計算を Excel の関数を補助的に用い，手作業で最大化するための手順を示す．次に，対数尤度関数をパラメータに関して偏微分した偏微分ベクトル，更に偏微分をした 2 階の偏微分行列（ヘッセ行列）の導出を丁寧に示す．それらを Excel シート上に展開し，ニュートン・ラフソン法による対数尤度の最大化の方法について示す．対数尤度の最大化の方法について，1 群の場合，回帰の場合，対数リンクでの回帰の場合，オフセットがある場合について第 1 章で取り上げた事例を用いて Excel を用いることにより計算過程を可視化しつつ示す．

2.1. 手作業による逐次的な対数尤度の最大化

表 1.4 の「有害雑草の種子の数」では，算術平均（ポアソン分布の位置パラメータ μ ）を用いて有害雑草の種子の数に対するポアソン分布の確率を計算した [スネデカーら (1972)]．本節では，算術平均を用いずに，探索的に対数尤度を最大化するような位置パラメータ $\hat{\mu}$ を変化させて推定する方法を提示する．理論的には自明であり，ナンセンスかもしれないが，一足飛びにポアソン回帰から始めるよりも，レンガを積むがごとく段階を追って理解を深めることは，応用力の根源となる．

表 2.1 に初期値を $\hat{\mu}=2.0$ とし，種子の数 y_i に対するポアソン分布の確率 P_i を Excel の `Poisson.dist()` 関数

$$\begin{aligned} P_i &= \frac{\mu^{y_i} e^{-\mu}}{y_i!} \\ &= \frac{2^{y_i} e^{-2}}{y_i!} = \text{Poisson.dist}(y_i, 2, \text{false}), \quad i=1, 2, \dots, 9 \end{aligned} \quad (2.1)$$

で計算した結果を示す．なお，引数の *false* は確率の計算のための引数で，*true* とすると下側確率のための引数となる．それぞれの種子の数 y_i に対する観測度数が n_i とすると，同時に起きる確率は $P_i^{n_i}$ （尤度 L_i ）となり，それらを全て掛け合わせた尤度 L は，

$$\begin{aligned} L &= \prod_{i=1}^9 L_i \\ &= \prod_{i=1}^9 P_i^{n_i} \\ &= 0.1353^3 \times 0.2707^{17} \times \cdots \times 0.0002^1 = 3.2034 \times 10^{-93} \end{aligned} \quad (2.2)$$

となり，その積は極めて小さく，数値計算に適していない．そこで，尤度の対数を取っても大小関係は保たれるので，対数尤度 $\ln L$ として扱うことが一般的である．

$$\begin{aligned} \ln L &= \sum_{i=1}^9 \ln L_i \\ &= \sum_{i=1}^9 \ln(P_i^{n_i}) \\ &= -6.0000 - 22.2165 - \cdots - 8.5635 = -212.9762 \end{aligned} \quad (2.3)$$

表 2.1 初期値 $\mu = 2.0$ に対する対数尤度の計算シート

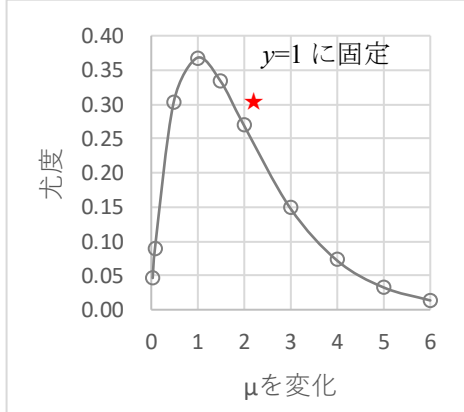
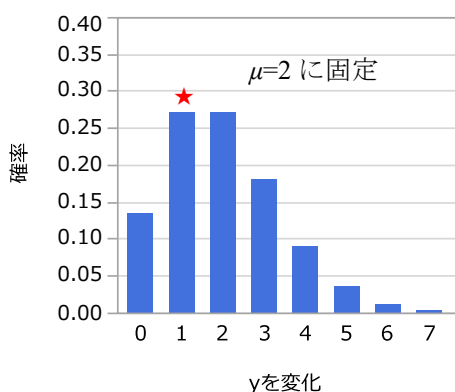
	$\mu^{\wedge} =$	2.0000	$\ln L =$	-212.9761953	
i	有害種子の数 y	観測度数 n	ポアソン確率 P	n 個のポアソン確率 P^n : 尤度 L_i	対数尤度 $\ln(P^n)$ $\ln L_i$
1	0	3	0.1353	2.4788E-03	-6.0000
2	1	17	0.2707	2.2465E-10	-22.2165
3	2	26	0.2707	1.7517E-15	-33.9782
4	3	16	0.1804	1.2636E-12	-27.3971
5	4	18	0.0902	1.5695E-19	-43.2984
6	5	9	0.0361	1.0385E-13	-29.8958
7	6	3	0.0120	1.7409E-06	-13.2611
8	7	5	0.0034	4.7968E-13	-28.3657
9	9	1	0.0002	1.9095E-04	-8.5635
	計	98	0.9991	3.2034E-93	-212.9762
				積	和

なぜ，確率ではなく“尤度”と言い換えるのか．表 2.2 左に示すように位置パラメータが $\mu = 2.0$ と定まっている場合に，ポアソン分布の確率関数を用いて観測データ y_i （正の整数）が起きる確率が計算されている．逆に，表 2.2 右に示すように観測データがある値 $y_i = 1$ と定まっている場合に，位置パラメータ μ （正の実数）を変化させたとき，ポアソン分布の確率関数を用いて $y_i = 1$ の確率を求めることができる．

この確率は， μ を連続的に変化させたときときに計算でき連続的に変化する．例えば， $y_i = 1$ としたときに， μ を 0.05 から 6 まで段階的に変化させた場合に，ポアソン分布の確率を計算ができる．ただし，これをポアソン分布の確率関数とは言えないので，“尤度関数”として別

表 2.2 確率関数と尤度関数の違い

μ を固定	y を変化	確率 P	μ を変化	y を固定	尤度 L_i	対数尤度
2	0	0.1353	0.05	1	0.0476	-3.0457
2	1	0.2707	0.1	1	0.0905	-2.4026
2	2	0.2707	0.5	1	0.3033	-1.1931
2	3	0.1804	1	1	0.3679	-1.0000
2	4	0.0902	1.5	1	0.3347	-1.0945
2	5	0.0361	2	1	0.2707	-1.3069
2	6	0.0120	3	1	0.1494	-1.9014
2	7	0.0034	4	1	0.0733	-2.6137
			5	1	0.0337	-3.3906
			6	1	0.0149	-4.2082



の呼び方が伝統的に使われている．そのため，尤度関数の考え方で計算されたポアソン分布の確率を“尤度”と区別する．なお，対数尤度は，自然対数が使われている．

対数尤度 $\ln L$ を最大化するようなポアソン分布の位置パラメータ $\hat{\mu}$ を求めたい．位置パラメータ μ は，算術平均となることが理論的に導出されているので，最尤法を適用する必要は全くない．しかし，ポアソン回帰の推定値は $\hat{\mu}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ であり， $\hat{\mu}_i$ の推定に算術平均を用いることはできない．一足飛びにポアソン回帰の場合の最尤法を扱う前に，結果が自明である一群の場合について最尤法の基本的な計算方法を示す．

最尤法は，表 2.3 に示すように，推定したい位置パラメータ $\hat{\mu}$ が未知であることを前提にし，対数尤度 $\ln L$ が最大になるように $\hat{\mu}$ を何らかの方法で逐次的に変化させ，変化の差分があらかじめ定めた極めて小さい値 δ よりも小さくなった場合の $\hat{\mu}$ を推定値とする方法である．

実際のデータの分布は，種子の数 $y_3 = 2$ よりも大きい方に裾を引いているので，表 2.1 に示すように $\hat{\mu} = 3.0$ と大きめに設定し対数尤度を計算すると $\ln L = -190.9585233$ となり，表 2.3 に転記して $\hat{\mu} = 2.0$ の $\ln L$ との差を求めると 22.0176720 の増加である．更に $\hat{\mu} = 3.1$ に増加すると

表 2.3 逐次的な挟み撃ち法による最尤解

$\hat{\mu}$	$\hat{\mu}$ 増減	対数尤度 $\ln L$	差分	
2.0		-212.9761953		
3.0	+	-190.9585233	22.0176720	
3.1	+	-191.0527358	-0.0942124	
3.05	-	-190.9658499	0.0868858	
3.02	-	-190.9517387	0.0141113	
3.025	+	-190.9520777	-0.0003390	
3.023	-	-190.9518449	0.0002328	
3.021	-	-190.9517416	0.0001032	
3.0205	-	-190.9517361	0.0000055	
3.0204	-	-190.9517360	0.0000001	最尤解
3.0203	+	-190.9517362	-0.0000002	

図 2.1 を参照のこと

$\ln L = -191.0527358$ となり -0.0942124 の減少となる. 従って $\ln L$ を最大にする $\hat{\mu}$ は $(3.0, 3.1)$ の範囲に存在することがわかる. このように $\ln L$ の差分を観察しつつ $\hat{\mu}$ を挟み撃ち的に適宜増減することにより最尤解として $\hat{\mu} = 3.0204$ が得られる.

表 2.1 の対数尤度の計算欄を右方向に伸ばし, 表 2.4 に示すように位置パラメータ $\hat{\mu}$ をある範囲に限定し, 対数尤度を計算しグラフを眺めながらその範囲を絞るような方法でも最尤解を得ることができる. 表頭に示すように $\hat{\mu}$ を, 2.950 から 0.025 刻みで 3.100 まで $j=1, 2, \dots, 7$ と 7 段階に変化させ, それぞれの y_i と度数 n_i について

$$\begin{aligned} \ln L_{ij} &= \ln[\text{Poisson.dist}(y_i, \hat{\mu}_{.j}, \text{false})^{n_i}] \\ \ln L_{.j} &= \sum_{i=1}^9 \ln L_{ij} \end{aligned} \quad (2.4)$$

表 2.4 対数尤度の最大化のための 0.025 刻みでの計算

		j	1	2	3	4	5	6	7
		$\hat{\mu}$	2.950	2.975	3.000	3.025	3.050	3.075	3.100
		$\ln L$	-191.0334	-190.9855	-190.9585	-190.9521	-190.9658	-190.9995	-191.0527
i	y	n	$\ln L_{.1}$	$\ln L_{.2}$	$\ln L_{.3}$	$\ln L_{.4}$	$\ln L_{.5}$	$\ln L_{.6}$	$\ln L_{.7}$
1	0	3	-8.8500	-8.9250	-9.0000	-9.0750	-9.1500	-9.2250	-9.3000
2	1	17	-31.7593	-32.0409	-32.3236	-32.6075	-32.8926	-33.1788	-33.4662
3	2	26	-38.4680	-38.6791	-38.8940	-39.1124	-39.3345	-39.5600	-39.7889
4	3	16	-23.9415	-23.9364	-23.9348	-23.9364	-23.9414	-23.9495	-23.9609
5	4	18	-32.4150	-32.2574	-32.1049	-31.9574	-31.8148	-31.6770	-31.5440
6	5	9	-20.9562	-20.8014	-20.6499	-20.5014	-20.3561	-20.2137	-20.0743
7	6	3	-9.1153	-9.0384	-8.9627	-8.8884	-8.8152	-8.7433	-8.6725
8	7	5	-19.5126	-19.3423	-19.1744	-19.0089	-18.8459	-18.6851	-18.5267
9	9	1	-6.0156	-5.9646	-5.9143	-5.8646	-5.8156	-5.7671	-5.7192

を用いて対数尤度 $\ln L_j$ を計算している．対数尤度 $\ln L_j$ が最も大きくなるのは， $j=4$ ， $\hat{\mu}_4=3.025$ ， $\ln L_4=-190.9521$ である．計算された対数尤度を図 2.1 左に示す．最大となるのは， $\hat{\mu}=3.025$ より少し小さいところにあることが推測される．

表 2.5 に示すように，更に刻み幅を 3.014 から 0.002 刻みで 3.026 まで変化させた結果を図 2.1 右に示す．この結果から，最大となるのは， $\hat{\mu}=3.020$ より少し大きいところにあると推測される．

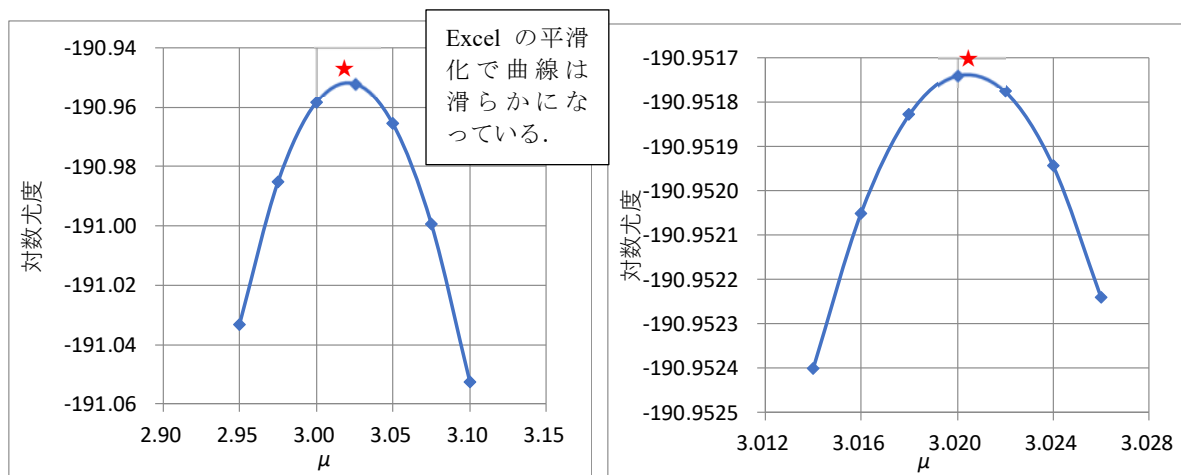


図 2.1 $\hat{\mu}$ を変化させた場合の対数尤度関数

表 2.5 対数尤度の最大化のための 0.002 刻みでの計算

		j	1	2	3	4	5	6	7
		母数 μ^{\wedge}	3.014	3.016	3.018	3.020	3.022	3.024	3.026
		$\ln L$	-190.9524	-190.9521	-190.9518	-190.9517	-190.9518	-190.9519	-190.9522
i	y	n	$\ln L_1$	$\ln L_2$	$\ln L_3$	$\ln L_4$	$\ln L_5$	$\ln L_6$	$\ln L_7$
1	0	3	-9.0420	-9.0480	-9.0540	-9.0600	-9.0660	-9.0720	-9.0780
2	1	17	-32.4824	-32.5052	-32.5279	-32.5506	-32.5734	-32.5961	-32.6189
3	2	26	-39.0159	-39.0334	-39.0509	-39.0685	-39.0860	-39.1036	-39.1213
4	3	16	-23.9353	-23.9354	-23.9356	-23.9358	-23.9360	-23.9363	-23.9366
5	4	18	-32.0217	-32.0099	-31.9982	-31.9865	-31.9748	-31.9632	-31.9516
6	5	9	-20.5664	-20.5545	-20.5427	-20.5309	-20.5191	-20.5073	-20.4956
7	6	3	-8.9209	-8.9150	-8.9091	-8.9031	-8.8972	-8.8913	-8.8854
8	7	5	-19.0814	-19.0682	-19.0550	-19.0418	-19.0286	-19.0155	-19.0024
9	9	1	-5.8864	-5.8824	-5.8785	-5.8745	-5.8706	-5.8666	-5.8627

2.2. Excel のソルバーによる対数尤度の最大化

統計ソフトを用いれば、対数尤度を全く意識することなくポアソン回帰の解を求めることができる。このことが、統計モデルの理論を学習し理解する意欲の妨げになっているのではないだろうか。統計ソフトから出力される結果に対して何らかの解釈をすることはできても、その理論的背景を関係者に説明し、納得してもらえることができるのだろうか。前節で示したように、Excel の計算機能を用いた逐次的な方法によって対数尤度を最大化する経験が、学習の第 1 歩であり、説明したい推定値を自ら計算し、解釈することができるようになる。

第 2 歩目は、Excel のソルバーを用いて、対数尤度を最大化する方法の習得である。Excel のソルバーは、手作業による逐次的な手順を代行してくれる。ただし、この Excel のソルバーの最適化の方法は、私にとっても多くの利用者にとってもブラック・ボックスである。ただし、対数尤度関数を利用者が、きちっと Excel で計算式として明示できることが必須であり、理論の学習にとって役に立つ。この方法により、統計ソフトがサポートしていない各種の統計モデルについて解析が迅速に行えるようになる。

Excel のソルバーと同様に、R 言語の `optim()` 関数も同様な最適化のための関数である。私は、Excel のソルバーを使う前は、SAS の IML 行列計算言語および JMP のスクリプト言語の `maximize()` 関数を使っていた。これらに比べ Excel の計算機能は比べようのないくらい劣っているが、計算過程および結果の可視化の面で優れている Excel のソルバーを多用するようになった。

第 3 歩目が、第 2.3 節に示す対数尤度関数の 2 階の偏微分行列を用いたニュートン・ラフソン法による最大化である。なお、一般化線形モデルに特徴的な、反復重み付き回帰による最尤法については、第 5 章で取り上げる。

表 2.6 左は初期値 $\hat{\mu} = 2.0$ に対する有害種子の数 y_i について、表 2.1 と同様に n_i 個のポアソン確率

$$L_i = P_i^{n_i} = \text{Poisson.dist}(y_i, 2, \text{false})^{n_i} \quad (2.5)$$

を計算し、対数を取って

$$\ln L_i = \ln P_i^{n_i} = \ln [\text{Poisson.dist}(y_i, 2, \text{false})^{n_i}] \quad (2.6)$$

とし、それぞれの i の対数尤度 $\ln L_i$ を計算している。それらの和

$$\begin{aligned} \ln L &= \sum_{i=1}^9 \ln L_i \\ &= \sum_{i=1}^9 \ln P_i^{n_i} = -6.0000 - 22.2165, \dots, -8.5635 = -212.9762 \end{aligned}$$

が、対数尤度 $\ln L = -212.9762$ である。

表 2.6 ソルバーを用いた対数尤度の最大化

			μ^{\wedge}	$\ln L$		μ^{\wedge}	$\ln L$
			2.0000	-212.9762		3.0204	-190.9517
i	y	n	p^n	$\ln(p^n)$		p^n	$\ln(p^n)$
1	0	3	2.4788E-03	-6.0000		1.1608E-04	-9.0612
2	1	17	2.2465E-10	-22.2165		7.2681E-15	-32.5553
3	2	26	1.7517E-15	-33.9782		1.0745E-17	-39.0721
4	3	16	1.2636E-12	-27.3971		4.0252E-11	-23.9359
5	4	18	1.5695E-19	-43.2984		1.2867E-14	-31.9841
6	5	9	1.0385E-13	-29.8958		1.2151E-09	-20.5285
7	6	3	1.7409E-06	-13.2611		1.3613E-04	-8.9019
8	7	5	4.7968E-13	-28.3657		5.3878E-09	-19.0391
9	9	1	1.9095E-04	-8.5635		2.8124E-03	-5.8737
	計	98	初期値(表2.1再掲)			ソルバーによる最大化	

表 2.6 右は、まず表 2.6 左を Excel シート上でコピーし、同じ結果が得られるように計算式のセル位置を調整する。次に、Excel の「データ」タブに含まれている「ソルバー」を起動し、対数尤度の数値セルを「目的セルの設定」ボックスに設定し、「目標値」が最大値に選択されていることを確認し、位置パラメータ μ のセルを「変数セルの変更」ボックスに設定し、「解決」ボタンのクリックで計算されたものである。結果として最適解の $\hat{\mu}=3.0204$ が得られており、対数尤度 $\ln L = -190.9517$ となっている。なお、「ソルバー」は、「分析ツール」と同様に Excel のアドインなので、使用できるような操作を前もって行う必要がある。

ソルバーのパラメーター

目的セルの設定:(I)		\$I\$7
目標値:	<input checked="" type="radio"/> 最大値(M)	<input type="radio"/> 最小値(N) <input type="radio"/> 指定値:(V)
変数セルの変更:(B)		\$H\$7

Excel のソルバーで、対数尤度を最大化しようとしても、設定した初期値によってはエラーが起きることもある。その場合は、対数尤度が大きくなるように手作業で初期値を変えて実行する必要がある。表 2.6 の例であれば、 $\hat{\mu}=8$ までは収束するが、 $\hat{\mu}=9$ にするとエラーとなる。 $\hat{\mu}=0$ とすると関数計算が不能となるが、 $\hat{\mu}=0.001$ とすれば、解が求まる。このような試行錯誤による結果が目の前で瞬時に確認できることは、Excel のソルバーによる最尤法の学習に欠かせない。

2.3. ニュートン・ラフソン法による対数尤度の最大化

ニュートン・ラフソン法は、ポアソン回帰などの一般化線形モデルのみならず打ち切りデータを含むワイブル回帰など、多くの統計モデルに対する汎用の解析方法として知られている。「最尤法とはどのような方法なのですか」との質問に、「尤度を最大化する方法です」とのまともではない回答がはびこる原因は、ニュートン・ラフソン法を用いて尤度を最大化するための具体的な数値例の例示がないためである。数値例を Web 上でも身近にある統計の教科書を探しているが見出すことができない。

ニュートン・ラフソン法については、Web 上で丁寧な説を見出すことができる。ただし、例示されているのは、2 次式の解を求めるなどの簡単な事例がほとんどである。ポアソン回帰を含む一般化線形モデルの場合には、第 1.4 節で示したように反復重み付き回帰による最尤法で対応可能であるが、より一般的な解法であるニュートン・ラフソン法をポアソン回帰で習得することは、打ち切りデータを含むワイブル回帰などへの応用力を身に付けることになる。詳しくは、高橋（2015）、「寿命試験データの統計解析」、高橋（2018）、「正規分布を仮定した打ち切りデータを含む回帰分析入門」を参照してもらいたい。

ニュートン・ラフソン法による最尤法の良さは、対数尤度を最大化する計算過程の中で、パラメータの共分散行列そのものを用いていることにある。共分散行列が得られれば、回帰直線の 95%信頼区間など各種の推定値に対する検定統計量の計算も容易にできる。Excel のソルバーで対数尤度関数を最大化することにより、統計モデルのパラメータの推定値を得ることができるが、各種の推論のためのパラメータの分散、パラメータ間の共分散が残念ながら得られない。そのために、ニュートン・ラフソン法によりパラメータの共分散行列を求める必要がある。

比較的容易なポアソン回帰について 2 階の偏微分行列（ヘッセ行列）を用いた最尤法で計算することができるようになれば、2 値データに対する一般化線形モデルに対しても対応できる力を付けることになる。なお、2 値データへの応用は、高橋（2017）「一般化線形モデルを Excel で極め活用するープロビット法・ロジット法・補 2 重対数法ー」を参照してもらいたい。

対数尤度関数の偏微分

ニュートン・ラフソン法を実施するためには、対数尤度関数をパラメータに関して 1 階の偏微分ベクトル \mathbf{U} (スコアベクトル)、さらに偏微分した 2 階の偏微分行列 \mathbf{H} (ヘッセ行列) を必要とする。一般的な統計の教科書で、偏微分が出てくるのは、回帰分析のパラメータ (β_0 , β_1) の計算式を求める際に、[第 4.2 節](#)で示すように誤差平方和 S_e を

$$S_e = \sum_i (y_i - \beta_0 - \beta_1 x_i)^2$$

パラメータ (β_0 , β_1) で偏微分して正規方程式を導出する場合に限られている。

ポアソン回帰の場合は、 $\mu_i = \beta_0 + \beta_1 x_i$ として、

$$\begin{aligned} \ln L &= \sum_i \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \\ &= \sum_i \ln \left[\frac{(\beta_0 + \beta_1 x_i)^{y_i} e^{-(\beta_0 + \beta_1 x_i)}}{y_i!} \right] \end{aligned} \quad (2.7)$$

をパラメータ (β_0 , β_1) で偏微分するのであるが、回帰分析の場合に比べて、かなり複雑である。

そこで、より簡単な 1 群のポアソン分布のあてはめについてニュートン・ラフソン法の適用を最初のステップとする。ポアソン分布の確率の対数を対数尤度 $\ln L_i$ とし、それらの合計を $\ln L$ と表記してきた。それぞれの y_i に対するポアソン分布の確率 P_i は、位置パラメータ μ に対して

$$P_i = \frac{\mu^{y_i} e^{-\mu}}{y_i!}, \quad i=1,2,\dots \quad (2.8)$$

なので、 y_i に対する μ を変化させた場合の確率を $L_i = P_i$ (尤度) と置き換える。対数尤度 $\ln L$ は、

$$\begin{aligned} \ln L &= \sum_i \ln L_i \\ &= \sum_i \ln \left(\frac{\mu^{y_i} e^{-\mu}}{y_i!} \right) \end{aligned} \quad (2.9)$$

となる。それぞれの y_i に、 n_i 個のデータがあるような場合の対数尤度 $\ln L_i$ は、

$$\begin{aligned} \ln L_i &= \ln \left(\frac{\mu^{y_i} e^{-\mu}}{y_i!} \right)^{n_i} \\ &= n_i (y_i \ln \mu - \mu - \ln y_i!) \end{aligned} \quad (2.10)$$

のように、 y_i のポアソン確率に対して n_i 乗とする。対数尤度関数 $\ln L_i$ をパラメータ μ で偏微分すると

$$\begin{aligned}
U_i &= \frac{\partial \ln L_i}{\partial \mu} \\
&= n_i \left(\frac{y_i}{\mu} - 1 \right) \\
&= n_i \frac{y_i - \mu}{\mu}
\end{aligned} \tag{2.11}$$

となり，さらにパラメータ μ で偏微分すると

$$\begin{aligned}
H_i &= \frac{\partial^2 \ln L_i}{\partial \mu^2} \\
&= n_i \frac{-y_i}{\mu^2}
\end{aligned} \tag{2.12}$$

が得られる．これらを i について加え 1 階の偏微分ベクトル \mathbf{U} と 2 階の偏微分行列 \mathbf{H} （ここでは共にスカラーであるが）を求める．

$$\mathbf{U} = \sum_i U_i = \sum_i \left(n_i \frac{y_i - \mu}{\mu} \right) \tag{2.13}$$

$$\mathbf{H} = \sum_i H_i = \sum_i \left(n_i \frac{-y_i}{\mu^2} \right) \tag{2.14}$$

数理統計の世界では，対数尤度関数の負の 2 階の偏微分行列を情報行列 \mathbf{I} （Fisher の情報量）というが，他の分野でも使われている一般的なヘッセ行列 $(-\mathbf{H})$ を使うことにする．また，対数尤度関数を数理統計の世界では「 l 」とするのが一般的であるが，尤度関数を L と表記し，対数尤度関数を $\ln L$ と区別して使うこととした．

反復計算

ニュートン・ラフソン法は，位置パラメータ μ の最初の初期値を $\hat{\mu}^{(0)}$ としたときに，

$$\hat{\mu}^{(1)} = \hat{\mu}^{(0)} + (-\mathbf{H}^{(0)})^{-1} \mathbf{U}^{(0)} \tag{2.15}$$

を繰り返し計算して，対数尤度 $\ln L$ を最大化する方法である．最大化したとの判断は，対数尤度の増分が， 10^{-6} 以下になった場合など適宜設定する．

ニュートン・ラフソン法は，対数尤度関数のパラメータで 1 階の偏微分したベクトル \mathbf{U} ，さらに 2 階の偏微分行列 \mathbf{H} を用いる方法として定式化されている．ポアソン回帰パラメータを $\hat{\boldsymbol{\beta}} = [\hat{\beta}_0 \ \hat{\beta}_1]^T$ としたときに，

$$\hat{\boldsymbol{\beta}}^{(1)} = \hat{\boldsymbol{\beta}}^{(0)} + (-\mathbf{H}^{(0)})^{-1} \mathbf{U}^{(0)} \tag{2.16}$$

と式 (2.15) と，全く同じ形式で回帰パラメータを推定することができる．

表 2.7 に示すように，(1) 回目の反復で初期値 $\hat{\mu}^{(0)} = 2.0$ をセットすると． $\hat{\mu}^{(1)} = 2.6757$ が計算される．計算過程は，順次， $(y_1 = 0, n_1 = 3)$ ， $(y_2 = 1, n_2 = 17)$ ， \dots ， $(y_9 = 9, n_9 = 1)$ と

し、最初の ($y_1=0$, $n_1=3$) については,

$$\text{初期値: } \hat{\mu}^{(0)} = 2.0$$

$$\text{ポアソン確率 } P_i: P_1^{(0)} = \text{Poisson.dist}(y_1, \mu^{(0)}, \text{false}) = 0.1353$$

$$\text{尤度 } L_i: L_1^{(0)} = (P_1^{(0)})^{n_1} = 0.1353^3 = 2.4788 \times 10^{-3}$$

$$\text{対数尤度 } \ln L_i: \ln L_1^{(0)} = \ln(2.4788 \times 10^{-3}) = -6.0000$$

$$1 \text{ 階の偏微分: } U_1^{(0)} = \frac{\partial \ln L_1}{\partial \mu} = n_1 \left(\frac{y_1 - \hat{\mu}^{(0)}}{\hat{\mu}^{(0)}} \right) = 3 \left(\frac{0-2}{2} \right) = -3.0000$$

$$2 \text{ 階の偏微分: } H_1^{(0)} = \frac{\partial^2 \ln L_1}{\partial \mu^2} = n_1 \frac{-y_1}{(\hat{\mu}^{(0)})^2} = 3 \frac{-0}{2^2} = 0.0000$$

と計算される。以下, $i=2,3,\dots,9$ について同様に Excel のフィルハンドルで計算式をコピーする。

表 2.7 ニュートン・ラフソン法 (初期値)

		($m-1$)	変化量	(m)	対数尤度	1階の	2階の	負の逆行列
		μ^{\wedge}	$(-H)^{-1}U$	μ^{\wedge}	$\ln L$	偏微分 U	偏微分 H	$(-H)^{-1}$
		2.0000	0.6757	2.6757	-212.9762	50.0000	-74.0000	0.0135
i	y	n	ポアソン P	L_i	$\ln L_i$	$/ \partial \mu$	$/ \partial \mu \partial \mu$	標準誤差
1	0	3	0.1353	2.4788E-03	-6.0000	-3.0000	0.0000	SE
2	1	17	0.2707	2.2465E-10	-22.2165	-8.5000	-4.2500	0.1162
3	2	26	0.2707	1.7517E-15	-33.9782	0.0000	-13.0000	
4	3	16	0.1804	1.2636E-12	-27.3971	8.0000	-12.0000	
5	4	18	0.0902	1.5695E-19	-43.2984	18.0000	-18.0000	
6	5	9	0.0361	1.0385E-13	-29.8958	13.5000	-11.2500	
7	6	3	0.0120	1.7409E-06	-13.2611	6.0000	-4.5000	
8	7	5	0.0034	4.7968E-13	-28.3657	12.5000	-8.7500	
9	9	1	0.0002	1.9095E-04	-8.5635	3.5000	-2.2500	
—— 表2.1 (再掲) ——								

次に、全ての i について以下の計算をする。

$$\text{対数尤度: } \ln L^{(0)} = \sum_{i=1}^9 \ln L_i^{(0)} = -6.0000 - 22.2165 \dots - 8.5635 = -212.9762$$

$$1 \text{ 階の偏微分: } U^{(0)} = \sum_{i=1}^9 U_i^{(0)} = -3.00 - 8.50 + \dots + 3.50 = 50.00$$

$$2 \text{ 階の偏微分: } H^{(0)} = \sum_{i=1}^9 H_i^{(0)} = 0.00 - 4.25 - \dots - 2.25 = -74.00$$

$$H \text{ の負の逆行列: } (-H^{(0)})^{-1} = -1 / (-74.00) = 0.0135$$

$$\begin{aligned} \hat{\mu}^{(1)} &= \hat{\mu}^{(0)} + (-H^{(0)})^{-1} U^{(0)} \\ \mu^{(1)} \text{ の計算: } &= 2.00 + 0.0135 \times 50.00 \\ &= 2.6757 \end{aligned}$$

$\hat{\mu}^{(1)}$ をコピーして、値のみを $\hat{\mu}^{(0)}$ にペースト

逐次計算： $\hat{\mu}^{(0)} \leftarrow \hat{\mu}^{(1)}$

すると、

表 2.8 に示すように以上の計算が第 2 回目として $\hat{\mu}^{(2)} = 2.9811$ が計算される。

さらに繰返しを行い、

表 2.8 に示すように 5 回目の反復で、 $\hat{\mu}^{(4)} = \hat{\mu}^{(5)} = 3.0240$ となったので、解が求まったと判定する。

表 2.8 ニュートン・ラフソン法による反復過程

反復(m)	$\mu^{(m-1)}$	$(-H)^{-1}U$	$\mu^{(m)}$	
1	2.0000	0.6757	2.6757	初期値
2	2.6757	0.3054	2.9811	
3	2.9811	0.0388	3.0199	
4	3.0199	0.0005	3.0204	
5	3.0204	0.0000	3.0204	収束

表 2.9 に 5 回目の反復で計算された結果を示す。表の右端が、負の偏微分 H の逆数 $(-H)^{-1} = 0.0308$ である。これが求めたかった推定値 $\hat{\mu}$ の分散の推定値である。平方根を取ると標準誤差 0.1756 が求まる。

表 2.9 ニュートン・ラフソン法による対数尤度の最大化

		(m-1)	変化量	(m)	対数尤度	1階の	2階の	負の逆行列
		μ^{\wedge}	$(-H)^{-1}U$	μ^{\wedge}	$\ln L$	偏微分 U	偏微分 H	$(-H)^{-1}$
		3.0204	0.0000	3.0204	-190.9517	0.0000	-32.4459	0.0308
i	y	n	ポアソン P	L_i	$\ln L_i$	$\partial \mu$	$\partial^2 \mu$	標準誤差
1	0	3	0.0488	1.1608E-04	-9.0612	-3.0000	0.0000	SE
2	1	17	0.1473	7.2682E-15	-32.5553	-11.3716	-1.8634	0.1756
3	2	26	0.2225	1.0745E-17	-39.0721	-8.7838	-5.7000	
4	3	16	0.2240	4.0252E-11	-23.9359	-0.1081	-5.2615	
5	4	18	0.1692	1.2867E-14	-31.9841	5.8378	-7.8923	
6	5	9	0.1022	1.2151E-09	-20.5285	5.8986	-4.9327	
7	6	3	0.0514	1.3613E-04	-8.9019	2.9595	-1.9731	
8	7	5	0.0222	5.3878E-09	-19.0391	6.5878	-3.8365	
9	9	1	0.0028	2.8124E-03	-5.8737	1.9797	-0.9865	

JMP による切片のみのポアソン回帰の適用

JMP の一般化線形モデルで、「リンク関数」を恒等、「分布」を Poisson, 「モデルの構成」は空白のまま（切片のみのモデル）とした結果を表 2.10 に示す。

表 2.10 JMP のポアソン回帰による推定結果

パラメータ推定値					推定値の共分散
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	共分散
切片	3.0204	0.1756	296	<.0001*	切片 切片 0.0308

元のデータ y_i の分散は、推定された $\hat{\mu}$ に等しいので、 $Var(y) = \hat{\mu} = 3.0204$ となるが、 $\hat{\mu}$ の標準誤差 SE は、0.1755 と推定されている。これは、表 2.9 で計算されていた負の 2 階の偏微分行列 H の逆数

$$H \text{ の負の逆行列: } (-H)^{-1} = 1 / 32.4459 = 0.0308,$$

$Var(\hat{\mu}) = 0.0308$ は、理論的なポアソン分布の分散は期待値 $\hat{\mu}$ に等しいことから、分散をデータ数で割って

$$Var(\hat{\mu}) = \frac{\hat{\mu}}{98} = \frac{3.0204}{98} = 0.0308$$

求めたものと等しくなり、その平方根は、 $SE(\hat{\mu}) = \sqrt{0.0308} = 0.1755$ として計算される。

一般的に、算術平均の SE は、不偏分散を自由度 ($N-1$) で割った平方根から得るのであるが、ポアソン分布を仮定した場合の SE とは異なる。なお、一般的な SE は、Excel の SumProduct() 関数を用いて $SE = 0.1829$ と計算される。

$N =$	98	=Sum(n の範囲)
平均 $\mu =$	3.0204	=SumProduct(y の範囲, n の範囲) / N
偏差平方和 $S =$	317.9592	=SumProduct(n の範囲, (y の範囲 - 平均) ^2)
分散 $V =$	3.2779	= $S / (N - 1)$
μ の $SE =$	0.1829	=Sqrt(V / N)

JMP による対数尤度関数の偏微分

ニュートン・ラフソン法の計算では、対数尤度関数をパラメータに関する偏微分式を正確に解いて、その数式を Excel シートに入力する必要がある。少しでも数式にミスがあると解が収束しない。偏微分式のプラス・マイナスに入れ間違いミスもよく経験する。そのために、JMP の偏微分の機能を使い、Excel シート上での偏微分式の計算結果が一致することを常に確認している。表 2.9 と同様な JMP シートを表 2.11 に示す。「対数尤度 $\ln L$ 」が「 \ln_L 」に、「1 階の偏微分 U 」が「 $d\mu$ 」に、「2 階の偏微分 H 」が「 $d\mu\mu$ 」に対応する。

「Poisson 確率 P 」には、計算式と式の引数 μ は、パラメータとして

計算式 : $\text{Poisson Probability}(\mu, y)$, パラメータ : $\mu = 3.0204$

が設定されていて、 y_i に対するポアソン分布の確率が計算され、表 2.9 の Excel と同じ結果と

表 2.11 対数尤度の JMP 計算式エディタを用いた偏微分

	i	y	n	Poisson 確率P	期待度数	ln_L	dμ	dμμ
1	0	0	3	0.0488	2.20	-9.0612	-3.0000	0.0000
2	1	1	17	0.1473	6.63	-32.5553	-11.3716	-1.8634
3	2	2	26	0.2225	10.01	-39.0721	-8.7838	-5.7000
4	3	3	16	0.2240	10.08	-23.9359	-0.1081	-5.2615
5	4	4	18	0.1692	7.61	-31.9841	5.8378	-7.8923
6	5	5	9	0.1022	4.60	-20.5285		
7	6	6	3	0.05				
8	7	7	5	0.02				
9				0.00			0.0000	0.0000
10				0.00			1.9797	-0.9865

なっている。対数尤度 \ln_L には、計算式

$$\ln_L : \ln L_i = n_i (y_i \ln \mu - \mu - \ln y_i!) = n \cdot \left(y \cdot \ln(\mu) - \mu - \ln(y!) \right)$$

が設定されている。偏微分 U の $d\mu$ には、

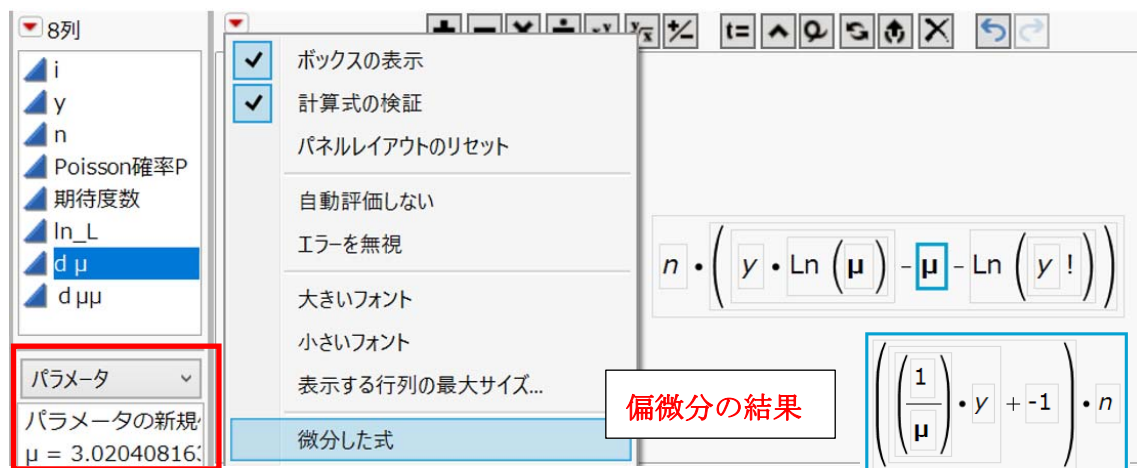
$$d\mu : U_i = \frac{n_i(y_i - \mu)}{\mu} = \left(\left(\frac{1}{\mu} \right) \cdot y + -1 \right) \cdot n$$

が計算式として設定されているが、表 2.12 に示すように、対数尤度 \ln_L の式を μ について、JMP の微分の機能を用いて微分した結果で、自ら計算式を入力したものではない。偏微分 $d\mu\mu$ は、対数尤度 \ln_L の式を μ について、2 回続けて微分した結果である。

$$d\mu\mu : H_i = n_i \frac{-y_i}{\mu^2} = -\mu^{-2} \cdot y \cdot n$$

表 2.11 の JMP シート上の数値と表 2.9 の Excel シートの数値が一致していれば、計算式の Excel シートへの入力ミスがないことが検証される。

表 2.12 JMP の計算エディタによる偏微分

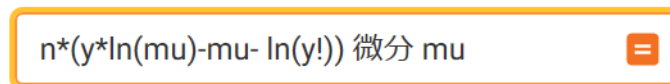


WolframAlpha による対数尤度関数の偏微分

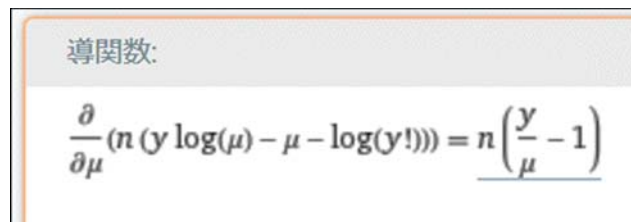
数学を支援するソフト無償版の WolframAlpha を用いて偏微分式を得ることもできる。

<https://ja.wolframalpha.com/> (2020 年 4 月 27 日アクセス)

計算したい式に対数尤度関数を入力すると



μ で微分した式が表示される。



JMP の偏微分の方法に比べ、洗礼された数式を得ることができる。このような数学ソフトの助けも借り、対数尤度関数の偏微分式を得ることにより、Excel による尤最法にチャレンジしてもらいたい。その経験が、最尤法は、どのような方法なのか」との質問に「尤度を最大にする方法である」との紋切り型な説明に引き続き、「実際には、...」との説明ができるようになることを願っている。

2.4. ポアソン回帰のバリエーション

ポアソン回帰には、いくつかのバリエーションがある．前章の第 1.4～1.13 節のタイトルを次に示す．

1.4.	人工データ（恒等リンク，3 水準，回帰）	16
1.5.	冠動脈心疾患の死亡者数（対数リンク，8 水準，オフセット，回帰）	23
1.6.	満月と新月の日の犯罪件数に対する尤度比検定（2 群）	27
1.7.	細菌を用いた試験データ（2×2 要因配置）	32
1.8.	細菌を用いた用量反応試験（恒等リンク，2 群，8 水準，効力比）	36
1.9.	植物の体サイズに関連した種子数（対数リンク，2 群，回帰）	40
1.10.	退役軍人における癌の発生（対数リンク，2 群，11 水準，オフセット）	46
1.11.	喫煙による冠動脈心疾患による死亡（対数リンク，2 群，5 水準，オフセット）	49
1.12.	医院への通院回数（過分散）	54
1.13.	雌のカブトガニに連結する雄の数（2 因子，2 変数，対数リンク，過分散）	56

仮定する分布は，全てポアソン分布であるが，（恒等リンク or 対数リンク），（オフセットなし or オフセットあり），（1 群 or 2 群 or 要因配置），（過分散なし，あり）などの組み合わせになっている．基本は，「オフセットなし」の「恒等リンク」でのポアソン回帰である．なお，第 2.5 節では，「オフセットなし」で「対数リンク」の場合，第 2.6 節では，「オフセットあり」で「対数リンク」の場合を取り上げる．過分散がある場合は，第 6 章で取り上げる

恒等リンクにおけるポアソン回帰の対数尤度

ポアソン回帰のモデル式は，ポアソン分布のパラメータ μ を回帰式に置き換えた式

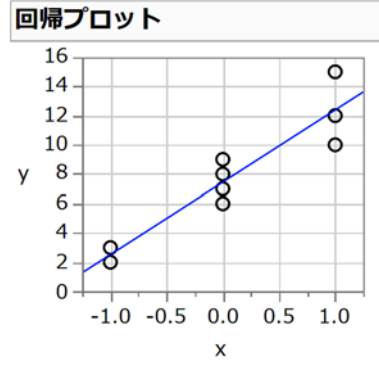
$$\mu_i = \beta_0 + \beta_1 x_i \quad i = 1, 2, \dots, n \quad (2.17)$$

$$\begin{aligned} P_i &= \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \\ &= \frac{(\beta_0 + \beta_1 x_i)^{y_i} e^{-(\beta_0 + \beta_1 x_i)}}{y_i!} \end{aligned} \quad (2.18)$$

である．第 1.4 節の人工データについてニュートン・ラフソン法によるポアソン回帰を行う．表 2.13 に表 1.7 のデザイン行列およびデータ y を再掲する [ドブソン (2008)]．

表 2.13 人工データに対するポアソン回帰のあてはめ (表 1.7 再掲)

i	x_0	x_1	y
1	1	-1	2
2	1	-1	3
3	1	0	6
4	1	0	7
5	1	0	8
6	1	0	9
7	1	1	10
8	1	1	12
9	1	1	15



それぞれの i について対数尤度 $\ln L_i$ は,

$$\begin{aligned}
 \ln L_i &= \ln \left[\frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \right] \\
 &= \ln \left[\frac{(\beta_0 + \beta_1 x_i)^{y_i} e^{-(\beta_0 + \beta_1 x_i)}}{y_i!} \right] \\
 &= y_i \ln(\beta_0 + \beta_1 x_i) - (\beta_0 + \beta_1 x_i) - \ln(y_i!)
 \end{aligned} \tag{2.19}$$

であり, 対数尤度 $\ln L_i$ をパラメータ β_0 および β_1 で偏微分すると

$$U_{1i} = \frac{\partial \ln L_i}{\partial \beta_0} = \frac{y_i}{\beta_0 + \beta_1 x_i} - 1 = \frac{y_i - \mu_i}{\mu_i} \tag{2.20}$$

$$U_{2i} = \frac{\partial \ln L_i}{\partial \beta_1} = \frac{\partial \ln L_i}{\partial \beta_0} x_i = \frac{y_i - \mu_i}{\mu_i} x_i \tag{2.21}$$

となり, さらに β_0 および β_1 で偏微分すると

$$\begin{aligned}
 H_{1,1,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} = \frac{-y_i}{(\beta_0 + \beta_1 x_i)^2} = \frac{-y_i}{\mu_i^2} \\
 H_{1,2,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_1} = \frac{\partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = \frac{-y_i}{\mu_i^2} x_i \\
 H_{2,1,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_0} = \frac{\partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = \frac{-y_i}{\mu_i^2} x_i \\
 H_{2,2,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_1} = \frac{\partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i^2 = \frac{-y_i}{\mu_i^2} x_i^2
 \end{aligned} \tag{2.22}$$

となる. これらを i について加えスコアベクトル \mathbf{U} とヘッセ行列 \mathbf{H} にまとめる.

$$\mathbf{U} = \begin{bmatrix} \sum_i U_{1i} \\ \sum_i U_{2i} \end{bmatrix} = \begin{bmatrix} \sum_i (y_i - \mu_i) / \mu_i \\ \sum_i x_i (y_i - \mu_i) / \mu_i \end{bmatrix} \tag{2.23}$$

$$\mathbf{H} = \begin{bmatrix} \sum_i H_{1,1,i} & \sum_i H_{1,2,i} \\ \sum_i H_{2,1,i} & \sum_i H_{2,2,i} \end{bmatrix} = \begin{bmatrix} \sum_i -y_i / \mu_i^2 & \sum_i -x_i y_i / \mu_i^2 \\ \sum_i -x_i y_i / \mu_i^2 & \sum_i -x_i^2 y_i / \mu_i^2 \end{bmatrix} \quad (2.24)$$

ニュートン・ラフソン法は，パラメータ β の最初の初期値を $\hat{\beta}^{(0)} = [\hat{\beta}_0^{(0)} \ \hat{\beta}_1^{(0)}]^T$ としたときに，次式により

$$\hat{\beta}^{(1)} = \hat{\beta}^{(0)} + (-\mathbf{H}^{(0)})^{-1} \mathbf{U}^{(0)} \quad (2.25)$$

繰り返し計算して，対数尤度 $\ln L$ を最大化する方法である．この式は，第 2.3 節でスカラーの場合の式 (2.15) と本質的に

$$\hat{\mu}^{(1)} = \hat{\mu}^{(0)} + (-\mathbf{H}^{(0)})^{-1} \mathbf{U}^{(0)} \quad (2.26)$$

同じである．最大化したとの判断は，対数尤度の増分が， 10^{-6} 以下になった場合など適宜設定する．

反復計算の実際

ニュートン・ラフソン法による計算方法は，第 2.3 節で位置パラメータが μ のみの場合について詳細に示した．ポアソン回帰の場合は， μ に変えて $\mu_i = \beta_0 + \beta_1 x_i$ のように 2 つのパラメータなので， 2×1 のスコアベクトル \mathbf{U} ， 2×2 のヘッセ行列 \mathbf{H} となるが，基本の計算原理は全く同じである．

表 2.14 に示すように初期値には，第 1.4 節の反復重み付き回帰の場合と同じ初期値

$$\hat{\beta}^{(0)} = \begin{bmatrix} 7.0 \\ 5.0 \end{bmatrix}$$

を用いて，偏微分式を用いた反復計算を行う．

表 2.14 初期値に対するニュートン・ラフソン法による計算結果

			元の(m-1)	変化量	新たな(m)	1階の	2階の偏微分		負の逆行列	
			パラメータ	$(-\mathbf{H})^{-1} \mathbf{U}$	パラメータ	偏微分 \mathbf{U}	\mathbf{H}		$(-\mathbf{H})^{-1}$	
		$\beta_0^{\wedge} =$	7.0000	0.4058	7.4058	0.8690	-2.1192	0.9931	0.6827	0.4499
		$\beta_1^{\wedge} =$	5.0000	-0.0091	4.9909	-0.4167	0.9931	-1.5069	0.4499	0.9601
i	x	y	μ^{\wedge}	P	$\ln L_i$	$\partial \beta_0$	$\partial \beta_1$	$\partial \beta_0 \partial \beta_0$	$\partial \beta_0 \partial \beta_1$	$\partial \beta_1 \partial \beta_1$
1	-1	2	2.0000	0.2707	-1.3069	0.0000	0.0000	-0.5000	0.5000	-0.5000
2	-1	3	2.0000	0.1804	-1.7123	0.5000	-0.5000	-0.7500	0.7500	-0.7500
3	0	6	7.0000	0.1490	-1.9038	-0.1429	0.0000	-0.1224	0.0000	0.0000
4	0	7	7.0000	0.1490	-1.9038	0.0000	0.0000	-0.1429	0.0000	0.0000
5	0	8	7.0000	0.1304	-2.0373	0.1429	0.0000	-0.1633	0.0000	0.0000
6	0	9	7.0000	0.1014	-2.2886	0.2857	0.0000	-0.1837	0.0000	0.0000
7	1	10	12.0000	0.1048	-2.2553	-0.1667	-0.1667	-0.0694	-0.0694	-0.0694
8	1	12	12.0000	0.1144	-2.1683	0.0000	0.0000	-0.0833	-0.0833	-0.0833
9	1	15	12.0000	0.0724	-2.6257	0.2500	0.2500	-0.1042	-0.1042	-0.1042
				計	-18.2021	0.8690	-0.4167	-2.1192	0.9931	-1.5069

最初の $i=1$ の場合は、次のように計算されている。

$$\hat{\mu}_1^{(0)} = \hat{\beta}_0^{(0)} + \hat{\beta}_1^{(0)} x_1 = 7.0 + 5.0 \times (-1) = 2.0$$

$$P_1^{(0)} = \text{Poisson.dist}(y_1, \hat{\mu}_1^{(0)}, \text{false}) = 0.2702$$

$$\ln L_1^{(0)} = \ln(p_1^{(0)}) = -1.3069$$

$$U_{1,i=1}^{(0)} = \frac{\partial \ln L_1}{\partial \hat{\beta}_0^{(0)}} = \frac{y_1 - \hat{\mu}_1^{(0)}}{\hat{\mu}_1^{(0)}} = \frac{2-2}{2} = 0$$

$$U_{2,i=1}^{(0)} = \frac{\partial \ln L_1}{\partial \hat{\beta}_1^{(0)}} = \frac{y_1 - \hat{\mu}_1^{(0)}}{\hat{\mu}_1^{(0)}} x_1 = \frac{2-2}{2} (-1) = 0$$

$$H_{1,1,i=1}^{(0)} = \frac{\partial \partial \ln L_1}{\partial \hat{\beta}_0^{(0)} \partial \hat{\beta}_0^{(0)}} = \frac{-y_1}{(\mu_1^{(0)})^2} = \frac{-2}{2^2} = -0.50$$

$$H_{1,2,i=1}^{(0)} = \frac{\partial \partial \ln L_1}{\partial \hat{\beta}_0^{(0)} \partial \hat{\beta}_1^{(0)}} x_1 = (-0.5) \times (-1) = 0.50$$

$$H_{2,1,i=1}^{(0)} = \frac{\partial \partial \ln L_1}{\partial \hat{\beta}_1^{(0)} \partial \hat{\beta}_0^{(0)}} x_1 = (-0.5) \times (-1) = 0.50$$

$$H_{2,2,i=1}^{(0)} = \frac{\partial \partial \ln L_1}{\partial \hat{\beta}_1^{(0)} \partial \hat{\beta}_1^{(0)}} x_1^2 = (-0.5) \times (-1)^2 = -0.50$$

すべての $i=1,2,\dots,9$ について逐次計算され、計の行に i について和が

	$\ln L_i$	$\partial \beta_0$	$\partial \beta_1$	$\partial \beta_0 \partial \beta_0$	$\partial \beta_0 \partial \beta_1$	$\partial \beta_1 \partial \beta_1$
計	-18.2021	0.8690	-0.4167	-2.1192	0.9931	-1.5069

計算されている。これらは、1 階の偏微分 U および 2 階の偏微分 H の欄に

1階の 偏微分 U	2階の偏微分 H	負の逆行列 $(-H)^{-1}$
0.8690	-2.1192 0.9931	0.6827 0.4499
-0.4167	0.9931 -1.5069	0.4499 0.9601

として代入されている。これらから、第 1 回目の反復は、 $\hat{\beta}^{(0)} = [7.0 \ 5.0]^T$ としたときに、

$$\begin{aligned}
\hat{\beta}^{(1)} &= \hat{\beta}^{(0)} + (-H^{(0)})^{-1} U^{(0)} \\
&= \begin{bmatrix} 7.0 \\ 5.0 \end{bmatrix} + \begin{bmatrix} 0.6827 & 0.4499 \\ 0.4499 & 0.9601 \end{bmatrix} \begin{bmatrix} 0.8690 \\ -0.4167 \end{bmatrix} \\
&= \begin{bmatrix} 7.0 \\ 5.0 \end{bmatrix} + \begin{bmatrix} 0.4058 \\ -0.0091 \end{bmatrix} \\
&= \begin{bmatrix} 7.4058 \\ 4.9909 \end{bmatrix}
\end{aligned}$$

による計算から、第 1 回目反復の $\hat{\beta}^{(1)} = [7.4057, 4.9909]^T$ が得られる。これをコピーして、 $\hat{\beta}^{(0)}$ に“値”のみをペーストすると、第 2 回目の反復計算が行われる。この繰返しを収束するまで続ける。少々面倒なので、ソルバーを用いて対数尤度 $\ln L$ の **-18.2021** の欄を最大化するように元の $(m-1)$ パラメータ ($\hat{\beta}_0$, $\hat{\beta}_1$) を変化させても良い。表 2.15 に示すように、第 4 回目の反復で対数尤度の増分も 0.0000 となり、パラメータの変化量も 0.0000 となり収束する。

表 2.15 ニュートン・ラフソン法による反復過程

反復		元の(m-1)	変化量	新たな(m)	負の逆行列		対数尤度	
		パラメータ	$(-H)^{-1}U$	パラメータ	$(-H)^{-1}$		$\ln L$	増分
1	$\beta_0^{\wedge} =$	7	0.4058	7.4058	0.6827	0.4499	-18.2021	
	$\beta_1^{\wedge} =$	5	-0.0091	4.9909	0.4499	0.9601		
2	$\beta_0^{\wedge} =$	7.4058	0.0445	7.4503			-18.0086	0.1934
	$\beta_1^{\wedge} =$	4.9909	-0.0532	4.9378				
3	$\beta_0^{\wedge} =$	7.4503	0.0013	7.4516			-18.0039	0.0047
	$\beta_1^{\wedge} =$	4.9378	-0.0025	4.9353				
4	$\beta_0^{\wedge} =$	7.4516	0.0000	7.4516	0.7817	0.4160	-18.0039	0.0000
	$\beta_1^{\wedge} =$	4.9353	0.0000	4.9353	0.4160	1.1915		

第 4 反復で収束した結果について行列 $(-H)^{-1}$ の対角要素について平方根をとり、推定されたパラメータの標準誤差 SE を推定する。

表 2.16 第 4 反復で収束した結果

	元の(m-1)	変化量	新たな(m)	1階の	2階の偏微分	負の逆行列			
	パラメータ	$(-H)^{-1}U$	パラメータ	偏微分U	H	$(-H)^{-1}$		SE	
$\beta_0^{\wedge} =$	7.4516	0.0000	7.4516	0.0000	-1.5711 0.5485	0.7817	0.4160	0.8842	
$\beta_1^{\wedge} =$	4.9353	0.0000	4.9353	0.0000	0.5485 -1.0308	0.4160	1.1915		1.0915

JMP によるポアソン回帰

表 2.17 に JMP の一般化線形モデルで、リンク関数を「恒等」、分布を「poisson」とした結果を示す。パラメータの推定値および標準誤差が一致していることが確認できる。もちろんパラメータの共分散行列，負のヘッセ行列の逆行列 $(-H)^{-1}$ と一致する。

表 2.17 JMP によるポアソン回帰の結果（表 1.7 再掲）

パラメータ推定値					推定値の共分散		
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	共分散		
切片	7.4516	0.8842	71.0299	<.0001*		切片	x
x	4.9353	1.0915	16.5260	<.0001*	切片	0.7817	0.4160
					x	0.4160	1.1915

複数の共変量をもつポアソン回帰の偏微分式

ここでは、単回帰の場合について示したが、共変量などを含む変数が複数ある場合への拡張は容易である。変数を増やし、 $\mathbf{x} = [1 \ x_{1i} \ x_{2i} \ \cdots \ x_{pi}]$ から成る $(p+1)$ 変数に（切片： $x_{0i}=1$ ）に拡張した場合、 $(p+1)$ のスコアベクトル $\mathbf{U}^{(p)}$ は、 \mathbf{x} に要素 $(y_i - \mu_i) / \mu_i$ を掛けた式

$$\mathbf{U}^{(p)} = \begin{bmatrix} \sum_i U_{0i} \\ \sum_i U_{1i} \\ \vdots \\ \sum_i U_{pi} \end{bmatrix} = \begin{bmatrix} \sum_i x_{0i} (y_i - \mu_i) / \mu_i \\ \sum_i x_{1i} (y_i - \mu_i) / \mu_i \\ \vdots \\ \sum_i x_{pi} (y_i - \mu_i) / \mu_i \end{bmatrix} \quad (2.27)$$

となり、 $(p+1) \times (p+1)$ のヘッセ行列 $\mathbf{H}^{(p)}$ は、2 次形式 $\mathbf{x}^T \mathbf{x}$ に要素 $[-y_i / \mu_i^2]$ を掛けた式

$$\mathbf{x}^T \mathbf{x} = [1 \ x_{1i} \ x_{2i} \ \cdots \ x_{pi}]^T [1 \ x_{1i} \ x_{2i} \ \cdots \ x_{pi}]$$

で求めることができる。

$$\begin{aligned} \mathbf{H}^{(p)} &= \begin{bmatrix} \sum_i H_{0,0,i} & \sum_i H_{0,1,i} & \cdots & \sum_i H_{0,p,i} \\ \sum_i H_{1,0,i} & \sum_i H_{1,1,i} & & \sum_i H_{1,p,i} \\ \vdots & & \ddots & \vdots \\ \sum_i H_{p,0,i} & \sum_i H_{p,1,i} & \cdots & \sum_i H_{p,p,i} \end{bmatrix} \\ &= \begin{bmatrix} \sum_i -y_i / \mu_i^2 & \sum_i -x_{1,i} y_i / \mu_i^2 & \cdots & \sum_i -x_{p,i} y_i / \mu_i^2 \\ \sum_i -3y_i / \mu_i^2 & \sum_i -x_{1,i}^2 y_i / \mu_i^2 & & \sum_i -x_{p,i} x_{1,i} y_i / \mu_i^2 \\ \vdots & & \ddots & \vdots \\ \sum_i -x_{p,i} y_i / \mu_i^2 & \sum_i -x_{1,i} x_{p,i} y_i / \mu_i^2 & \cdots & \sum_i -x_{p,i}^2 y_i / \mu_i^2 \end{bmatrix} \end{aligned} \quad (2.28)$$

2.5. 対数リンクの場合のポアソン回帰

対数リンク

変数 x_i が増大するにつれ、観測データ y_i が指数関数的に増加することもしばしば経験する。位置パラメータ μ_i も指数関数的に増加すると仮定すると、回帰式 $\beta_0 + \beta_1 x_i$ は、

$$\mu_i = \exp(\beta_0 + \beta_1 x_i) + \varepsilon_i, \quad \varepsilon_i \sim \text{ポアソン分布} \quad i = 1, 2, \dots, n \quad (2.29)$$

と定義されるのでポアソン確率 P_i は、

$$\begin{aligned} P_i &= \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \\ &= \frac{\exp(\beta_0 + \beta_1 x_i)^{y_i} e^{-\exp(\beta_0 + \beta_1 x_i)}}{y_i!} \end{aligned} \quad (2.30)$$

となる。推定式の形にして両辺に対数をとって、

$$\ln(\hat{\mu}_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i, \quad i = 1, 2, \dots, n \quad (2.31)$$

のように線形化することができる。一般化線形モデルでは、この線形化する変換をリンク関数と言う。一般化線形モデルに対する統計ソフトでは、元の指数関数ではなく、線形化する関数名を使っていて、この場合のリンク関数は「対数」である。2 値反応が 2 項分布に従うとした場合のリンク関数としては、(プロビット・ロジット・補対数-対数) などがある。

第 1.9 節「植物の体サイズに関連した種子数 (対数リンク, 2 群, 回帰)」で (対数リンク) としたのは、対数を取った場合に線形化できる事例とした [久保 (2012)]。図表は、元のデータのスケールを用いており、対数表示をした図表は示さなかった。表 1.35 から得られるパラメータを用いた回帰式は、対数変換した

$$\ln(\hat{\mu}_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i = 0.7459 + 0.1323 x_i$$

であるが、グラフでは元のスケールで、

$$\hat{\mu}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i) = \exp(0.7459 + 0.1323 x_i)$$

で表した。自然対数での目盛りは、元のスケールへの目視による換算は厄介なので、常用対数目盛とした結果を図 2.2 右に示す。切片は、 $\exp(0.7459) = 2.11$ となる。

この図 2.2 のポアソン回帰の 95%信頼区間は、JMP のオプションでファイルの出力した結果を、「重ね合わせプロット」で別途作図したものである。

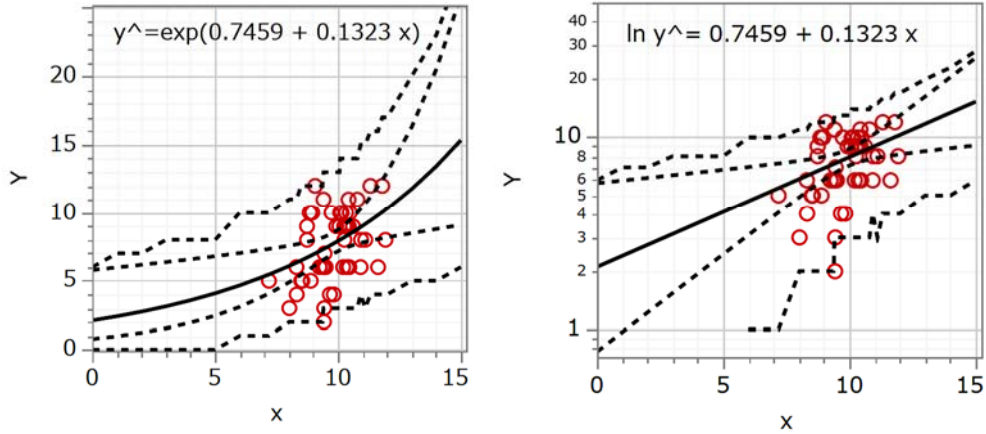


図 2.2 C 群に対するポアソン回帰のあてはめと 95%信頼区間の表示（表 1.35 再掲）

一般化線形モデルの解法としての反復重み付き回帰では、図 2.2 右の対数変換した線形式を用いて、図 2.2 左の指数関数に対するポアソン回帰を間接的にを行っている。詳細は、[第 5 章](#)を参照のこと。

対数リンクの場合の偏微分式

ニュートン・ラフソン法による最尤法は、図 2.2 左に示す指数関数に対して直接計算するので、図 2.2 左の対数変換した回帰式を使う必要はない。

$$\mu_i = \exp(\beta_0 + \beta_1 x_i) + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.32)$$

$$\begin{aligned} P_i &= \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \\ &= \frac{\exp(\beta_0 + \beta_1 x_i)^{y_i} e^{-\exp(\beta_0 + \beta_1 x_i)}}{y_i!} \end{aligned} \quad (2.33)$$

である。それぞれの i に対する対数尤度 $\ln L_i$ は、

$$\begin{aligned} \ln L_i &= \ln \left[\frac{\exp(\beta_0 + \beta_1 x_i)^{y_i} e^{-\exp(\beta_0 + \beta_1 x_i)}}{y_i!} \right] \\ &= y_i(\beta_0 + \beta_1 x_i) - \exp(\beta_0 + \beta_1 x_i) - \ln(y_i!) \end{aligned} \quad (2.34)$$

パラメータ β_0 および β_1 で偏微分すると

$$\begin{aligned} U_{1i} &= \frac{\partial \ln L_i}{\partial \beta_0} = y_i - \exp(\beta_0 + \beta_1 x_i) = y_i - \mu_i \\ U_{2i} &= \frac{\partial \ln L_i}{\partial \beta_1} = y_i x_i - \exp(\beta_0 + \beta_1 x_i) x_i = (y_i - \mu_i) x_i \end{aligned} \quad (2.35)$$

となり，さらに β_0 および β_1 で偏微分すると

$$\begin{aligned}
H_{1,1,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} = -\exp(\beta_0 + \beta_1 x_i) = -\mu_i \\
H_{1,2,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_1} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = -\mu_i x_i \\
H_{2,1,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_0} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = -\mu_i x_i \\
H_{2,2,i} &= \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_1} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i^2 = -\mu_i x_i^2
\end{aligned} \tag{2.36}$$

となる．これらを i について加えスコアベクトル \mathbf{U} とヘッセ行列 \mathbf{H} にまとめる．

$$\mathbf{U} = \begin{bmatrix} \sum_i U_{1i} \\ \sum_i U_{2i} \end{bmatrix} = \begin{bmatrix} \sum_i (y_i - \mu_i) \\ \sum_i (y_i - \mu_i) x_i \end{bmatrix} \tag{2.37}$$

$$\mathbf{H} = \begin{bmatrix} \sum_i H_{1,1,i} & \sum_i H_{1,2,i} \\ \sum_i H_{2,1,i} & \sum_i H_{2,2,i} \end{bmatrix} = \begin{bmatrix} \sum_i -\mu_i & \sum_i -\mu_i x_i \\ \sum_i -\mu_i x_i & \sum_i -\mu_i x_i^2 \end{bmatrix} \tag{2.38}$$

ニュートン・ラフソン法は，これまでも示してきたようにパラメータ $\boldsymbol{\beta} = [\hat{\beta}_0 \ \hat{\beta}_1]$ の最初の初期値を $\hat{\boldsymbol{\beta}}^{(0)}$ としたときに，

$$\hat{\boldsymbol{\beta}}^{(1)} = \hat{\boldsymbol{\beta}}^{(0)} + (-\mathbf{H}^{(0)})^{-1} \mathbf{U}^{(0)} \tag{2.39}$$

を繰り返し計算して，対数尤度 $\ln L$ を最大化する方法である．最大化したとの判断は，対数尤度の増分が， 10^{-6} 以下になった場合など適宜設定する．式 (2.36) は，式 (2.23) と形式は全く同じである．対数尤度が異なれば，スコアベクトルおよびヘッセ行列も異なるが，対数尤度を最大化する方式は，全く同じとなることが，ニュートン・ラフソン法の見通しの良さである．

Excel による反復計算

表 2.18 に Excel によるニュートン・ラフソン法で，対数尤度を最大化した結果を示す．計算シートの構成は，表 2.15 と同じであるが，埋め込まれている式が，「対数リンク」を反映している．異なるのは， $\hat{\mu}_i = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)$ ，偏微分式 $\partial \ln L_i / \partial \beta_0, \dots, \partial \partial \ln L_i / \partial \beta_0 \partial \beta_1$ の計算である．初期値として， $\hat{\beta}_0 = 0.70$ ， $\hat{\beta}_1 = 0.10$ を与えて，表 2.18 で示した反復の方法で，第 4 反復で収束した結果である．

表 2.18 C 群に対するニュートン・ラフソン法によるポアソン回帰のあてはめ

			元の(m-1) パラメータ	変化量 $(-H)^{-1}U$	新たな(m) パラメータ	1階の 偏微分U	2階の偏微分 H		負の逆行列 $(-H)^{-1}$	
		$\beta_0^{\wedge} =$	0.7459	0.0000	0.7459	0.00	-389.0	-3865.2	0.2657	-0.0265
		$\beta_1^{\wedge} =$	0.1323	0.0000	0.1323	-0.02	-3865.2	-38781	-0.0265	0.0027
i	x	y	μ^{\wedge}	p	$\ln Li$	$\partial \beta_0$	$\partial \beta_1$	$\partial \beta_0 \partial \beta_0$	$\partial \beta_0 \partial \beta_1$	$\partial \beta_1 \partial \beta_1$
1	8.31	6	6.3279	0.1592	-1.8373	-0.3279	-2.7245	-6.33	-52.58	-437.0
2	9.44	6	7.3479	0.1408	-1.9607	-1.3479	-12.7243	-7.35	-69.36	-654.8
3	9.50	6	7.4065	0.1392	-1.9716	-1.4065	-13.3613	-7.41	-70.36	-668.4
4	9.07	12	6.9970	0.0263	-3.6384	5.0030	45.3773	-7.00	-63.46	-575.6
5	10.16	10	8.0820	0.1013	-2.2900	1.9180	19.4865	-8.08	-82.11	-834.3
:										
49	11.32	12	9.4222	0.0827	-2.4926	2.5778	29.1811	-9.42	-106.66	-1207.4
50	9.25	6	7.1656	0.1453	-1.9291	-1.1656	-10.7815	-7.17	-66.28	-613.1
				計	-116.27	0.00	-0.02	-389.00	-3865.20	-38781

表 2.19 ニュートン・ラフソン法の反復

		元の(m-1) パラメータ	変化量 $(-H)^{-1}U$	新たな(m) パラメータ
1	$\beta_0^{\wedge} =$	0.7000	-0.0179	0.6821
	$\beta_1^{\wedge} =$	0.1000	0.0464	0.1464
2	$\beta_0^{\wedge} =$	0.6821	0.0565	0.7386
	$\beta_1^{\wedge} =$	0.1464	-0.0131	0.1333
3	$\beta_0^{\wedge} =$	0.7386	0.0073	0.7459
	$\beta_1^{\wedge} =$	0.1333	-0.0010	0.1323
4	$\beta_0^{\wedge} =$	0.7459	0.0000	0.7459
	$\beta_1^{\wedge} =$	0.1323	0.0000	0.1323

JMP による対数リンクでのポアソン回帰

JMP の一般化線形モデルによるポアソン回帰を行い、推定値および共分散行列を表 2.20 に示す。Excel での計算結果と一致していることが確認できる。

表 2.20 JMP による対数リンクでのポアソン回帰

パラメータ推定値					推定値の共分散		
項	推定値	標準誤差	尤度比カイ2乗	p値(Prob>ChiSq)	共分散		
切片	0.7459	0.5154	2.0730	0.1499		切片	x
x	0.1323	0.0516	6.5990	0.0102*	切片	0.2657	-0.026
					x	-0.026	0.0027

2.6. 対数リンクでオフセットがある場合のポアソン回帰

第 1.5 節の「冠動脈心疾患の死亡者数」のデータには、年齢階層別の死亡者数をカウントした結果である [アーミテジ (2001)]. そのため、人口統計で母集団の人数を知ることができ、死亡率を計算することが可能である. もちろん 2 項分布を仮定したロジスティック回帰を適用することも可能ではある. しかし、1%以下を対象とした推定となるので結果の解釈が煩わしい. そのために、一般的には、対数リンクを仮定し、分母を考慮するポアソン回帰の適用が好まれる.

表 2.21 オーストラリアのある地方の冠動脈心疾患の死亡者数 (表 1.11 再掲)

	年齢層	死亡者数	母集団	死亡率	10万比
i	x	y	人数 n		人数
1	30	1	17,742	0.006%	5.6
2	35	5	16,554	0.030%	30.2
3	40	5	16,059	0.031%	31.1
4	45	12	13,083	0.092%	91.7
5	50	25	10,784	0.232%	231.8
6	55	38	9,645	0.394%	394.0
7	60	54	10,706	0.504%	504.4
8	65	65	9,933	0.654%	654.4
	合計	205	104,506	0.196%	196.2

モデルは、死亡数が指数的に増えているので、母集団の人数 n_i を含んだ指数関数

$$\mu_i = n_i \exp(\beta_0 + \beta_1 x_i) + \varepsilon_i \quad i = 1, 2, \dots, n \quad (2.40)$$

で死亡数の推定をする. 推定式の形にして両辺対数を取ると

$$\ln(\mu_i) = \ln(n_i) + \beta_0 + \beta_1 x_i \quad i = 1, 2, \dots, n \quad (2.41)$$

オフセット $\ln(n_i)$ を含む線形式となる. これは、パラメータ β_0 と β_1 を共通とし、切片をそれぞれ $\ln(n_i)$ 分かさ上げをした回帰直線を求めることと解される. 母集団の人数を 10,000 人と固定すれば、 $\ln(10,000) = 9.2103$ をかさ上げすることに対応する.

対数を取らない式では、

$$P_i = n_i \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!} \quad (2.42)$$

$$= \frac{[n_i \exp(\beta_0 + \beta_1 x_i)]^{y_i} e^{-n_i \exp(\beta_0 + \beta_1 x_i)}}{y_i!}$$

である. ニュートン・ラフソン法による最尤法では、この式の対数を取った対数尤度関数をパラメータで偏微分して対応する.

それぞれの i に対する対数尤度 $\ln L_i$ は,

$$\ln L_i = \ln \left\{ \frac{[n_i \exp(\beta_0 + \beta_1 x_i)]^{y_i} e^{-n_i \exp(\beta_0 + \beta_1 x_i)}}{y_i!} \right\} \quad (2.43)$$

$$= y_i [\ln(n_i) + (\beta_0 + \beta_1 x_i)] - n_i \exp(\beta_0 + \beta_1 x_i) - \ln(y_i!)$$

であり, パラメータ β_0 および β_1 で偏微分すると

$$U_{1i} = \frac{\partial \ln L_i}{\partial \beta_0} = y_i - n_i \exp(\beta_0 + \beta_1 x_i) = y_i - \mu_i$$

$$U_{2i} = \frac{\partial \ln L_i}{\partial \beta_1} = y_i x_i - n_i \exp(\beta_0 + \beta_1 x_i) x_i = (y_i - \mu_i) x_i \quad (2.44)$$

$$\text{ただし, } n_i \exp(\beta_0 + \beta_1 x_i) = \mu_i$$

となり, さらに β_0 および β_1 で偏微分すると

$$H_{1,1,i} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} = -n_i \exp(\beta_0 + \beta_1 x_i) = -\mu_i$$

$$H_{1,2,i} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_1} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = -\mu_i x_i$$

$$H_{2,1,i} = \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_0} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i = -\mu_i x_i \quad (2.45)$$

$$H_{2,2,i} = \frac{\partial \partial \ln L_i}{\partial \beta_1 \partial \beta_1} = \frac{\partial \partial \ln L_i}{\partial \beta_0 \partial \beta_0} x_i^2 = -\mu_i x_i^2$$

となる. これらを i について加えスコアベクトル \mathbf{U} とヘッセ行列 \mathbf{H} にまとめる.

$$\mathbf{U} = \begin{bmatrix} \sum_i U_{1i} \\ \sum_i U_{2i} \end{bmatrix} = \begin{bmatrix} \sum_i (y_i - \mu_i) \\ \sum_i (y_i - \mu_i) x_i \end{bmatrix} \quad (2.46)$$

$$\mathbf{H} = \begin{bmatrix} \sum_i H_{1,1,i} & \sum_i H_{1,2,i} \\ \sum_i H_{2,1,i} & \sum_i H_{2,2,i} \end{bmatrix} = \begin{bmatrix} \sum_i -\mu_i & \sum_i -\mu_i x_i \\ \sum_i -\mu_i x_i & \sum_i -\mu_i x_i^2 \end{bmatrix} \quad (2.47)$$

なお, オフセットがない場合は, $\mu_i = \exp(\beta_0 + \beta_1 x_i)$ と n_i を含まないだけで, スコアベクトル \mathbf{U} とヘッセ行列 \mathbf{H} の形式は同じである.

ニュートン・ラフソン法は, パラメータ $\boldsymbol{\beta} = [\beta_0, \beta_1]^T$ の最初の初期値を $\hat{\boldsymbol{\beta}}^{(0)}$ としたときに,

$$\hat{\boldsymbol{\beta}}^{(1)} = \hat{\boldsymbol{\beta}}^{(0)} + (-\mathbf{H})^{-1} \mathbf{U} \quad (2.48)$$

を繰り返し計算して, 対数尤度 $\ln L$ を最大化する方法である. 最大化したとの判断は, 対数尤度の増分が, 10^{-6} 以下になった場合など適宜設定する.

表 2.22 にニュートン・ラフソン法による繰返し計算の Excel シートを示す．得られた推定値は， $\beta_0 = -11.6278$ ， $\beta_1 = 0.1044$ である．これは，表 1.11 の JMP でのポアソン回帰で求めた推定値に一致する．

表 2.22 冠動脈心疾患の死亡者数のニュートン・ラフソン法による解

				元の パラメータ	変化量 $(-H)^{-1}U$	新たな パラメータ	1階の 偏微分 U	2階の偏微分 H		負の逆行列 $(-H)^{-1}$	
				$\beta_0 =$	-11.6278	0.0000	-11.6278	0.0000	-205.0 -11750.0	0.2053	-0.0035
				$\beta_1 =$	0.1044	0.0000	0.1044	0.0000	-11750.0 -689869	-0.0035	0.0001
i	x	y	n	μ^{\wedge}	ポアソン P	$\ln Li$	$\partial \beta_0$	$\partial \beta_1$	$\partial \beta_0 \partial \beta_0$	$\partial \beta_0 \partial \beta_1$	$\partial \beta_1 \partial \beta_1$
1	30	1	17,742	3.629	0.0963	-41519	-2.63	-78.88	-3.63	-109	-3266
2	35	5	16,554	5.708	0.1676	-29568	-0.71	-24.78	-5.71	-200	-6992
3	40	5	16,059	9.335	0.0522	-47429	-4.33	-173.38	-9.33	-373	-14935
4	45	12	13,083	12.819	0.1113	-28718	-0.82	-36.87	-12.82	-577	-25959
5	50	25	10,784	17.812	0.0220	-41180	7.19	359.38	-17.81	-891	-44531
6	55	38	9,645	26.855	0.0083	-46161	11.14	612.97	-26.86	-1477	-81237
7	60	54	10,706	50.250	0.0473	-32670	3.75	224.99	-50.25	-3015	-180900
8	65	65	9,933	78.591	0.0142	-42284	-13.59	-883.43	-78.59	-5108	-332048
				計		-309529	0.00	0.00	-205	-11750	-689869

推定された回帰式は，

$$\begin{aligned}\hat{\mu}_i &= n_i \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i) \\ &= n_i \exp(-11.6278 + 0.1044 x_i)\end{aligned}$$

であり，両辺対数を取ると

$$\begin{aligned}\ln(\mu_i) &= \ln(n_i) + \beta_0 + \beta_1 x_i \\ &= \ln(n_i) - 11.6278 + 0.1044 x_i\end{aligned}$$

のようにオフセット $\ln(n_i)$ を含む線形式となる．推定された回帰式には，オフセット項を含んだ形になっていて，通常の回帰分析の結果と異なる，どのように理解したら良いのだろうか．オフセット $\ln(n_i)$ に実際の数値を入れた場合に，

$$\begin{aligned}\ln(\mu_i) &= \ln(17,742) - 11.6278 + 0.1044 x_i \\ 30 \text{ 歳 : } &= 9.7837 - 11.6278 + 0.1044 x \\ &= -1.8442 + 0.1044 x \\ &: \\ 65 \text{ 歳 : } &= \ln(9,933) - 11.6278 + 0.1044 x_i \\ &= -2.4242 + 0.1044 x\end{aligned}$$

のように，各年代の分母の大きさに応じて切片を上下させている．各年代について切片を計算した結果を表 2.23 に示す．

表 2.23 オフセット値を考慮した切片

	年齢層	死亡者数		母集団	オフセット	調整後	傾き
i	x	y	$\ln y$	人数 n	$\ln n$	$\ln n + \beta_0^\wedge$	β_1^\wedge
1	30	1	0.0000	17,742	9.7837	-1.8441	0.1044
2	35	5	1.6094	16,554	9.7144	-1.9134	0.1044
3	40	5	1.6094	16,059	9.6840	-1.9438	0.1044
4	45	12	2.4849	13,083	9.4791	-2.1487	0.1044
5	50	25	3.2189	10,784	9.2858	-2.3420	0.1044
6	55	38	3.6376	9,645	9.1742	-2.4536	0.1044
7	60	54	3.9890	10,706	9.2786	-2.3492	0.1044
8	65	65	4.1744	9,933	9.2036	-2.4242	0.1044
	合計	205		104,506	$\beta_0^\wedge =$	-11.6278	

図 2.3 に年齢と死亡者数の対数についての散布図上に、各年代別の回帰直線を重ね書きした結果を示す。直線は、切片まで伸ばさずに各年齢の前後に限定して描いた。この例では、母集団の人数に大きな差がなかったため、直線の上下の振れはわずかである。

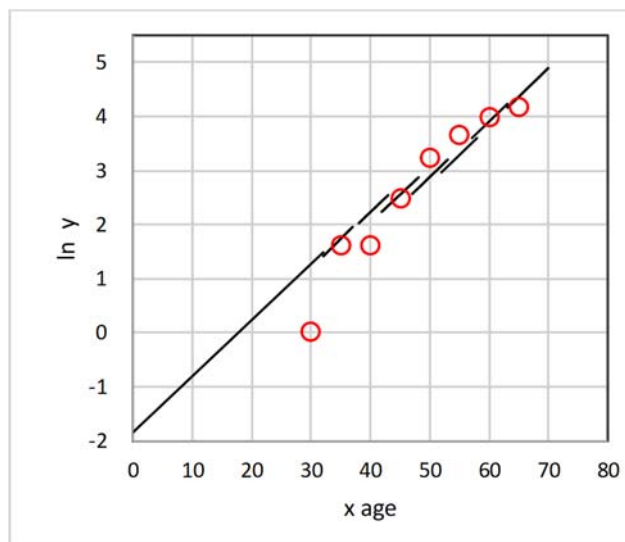


図 2.3 オフセットを含む場合の回帰直線の例示

オフセットがある場合には、母集団の平均的で切れの良い母集団の大きさ、この例であれば 10,000 人と設定し、 $x_1 = 30$ に対し 1 万人あたりの死亡数を計算し、

$$y_1^{(1万)} = 10,000 \frac{x_1}{n_1} = 10,000 \times \frac{1}{17,742} = 0.56$$

推定死亡者数を

$$\begin{aligned} \hat{\mu}_1^{(1万)} &= n_i \exp(\hat{\beta}_0 + \hat{\beta}_1 x_1) \\ &= 10,000 \times \exp(-11.6278 + 0.1044 \times 30) = 2.05 \end{aligned}$$

のように計算し，グラフ化することが望ましい。

表 2.24 1 万人比に換算したポアソン回帰の推定値

	年齢層	死亡者数	母集団	死亡者数	推定
i	x	y	n	1 万人比	1 万人比
1	30	1	17,742	0.56	2.05
2	35	5	16,554	3.02	3.45
3	40	5	16,059	3.11	5.81
4	45	12	13,083	9.17	9.80
5	50	25	10,784	23.18	16.52
6	55	38	9,645	39.40	27.84
7	60	54	10,706	50.44	46.94
8	65	65	9,933	65.44	79.12
	$\beta_0^{\wedge} =$	-11.6278	$\beta_1^{\wedge} =$	0.1044	

死亡者数 1 万人比および推定値について図 2.4 に示す。

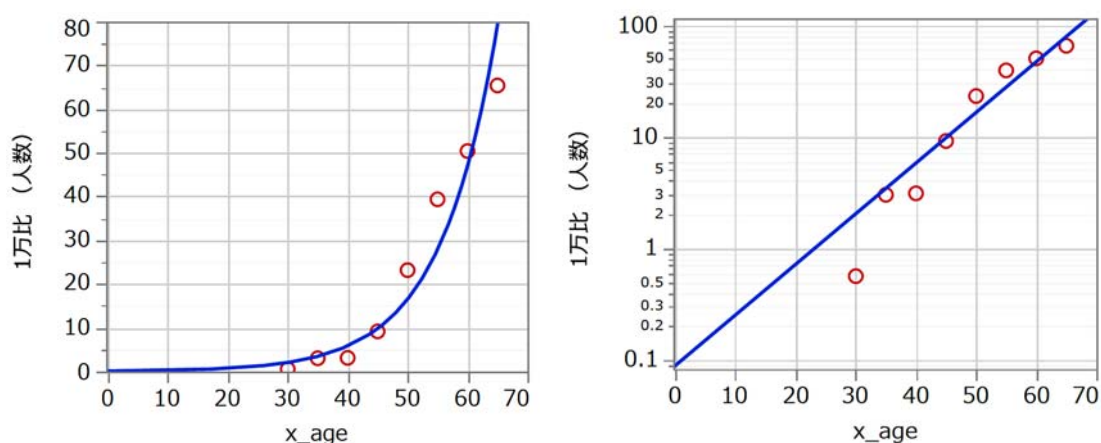


図 2.4 1 万人比で換算したポアソン回帰（曲線，直線）

オフセットがある場合のポアソン回帰で推定されたパラメータは，オフセットが 1 人の場合の死亡者数（死亡率）を計算していることになる。なお，第 5.5 節では，反復重み付き回帰によるポアソン回帰で 2 次式のあてはめた結果を示す。

2 値反応としたロジスティック回帰

表 1.14 に示したように「分布」を「2 項」，「リンク関数」を「ロジット」とした場合の一般線形モデルの結果から，

$$\beta_0 = -11.6395, \quad \beta_1 = 0.1047$$

が得られることを示した。これは，いわゆるロジスティック回帰であるが，表 2.25 に示すように，死亡率 π_i に対してロジット変換

$$\text{logit}_i = \ln \left(\frac{\pi_i}{1 - \pi_i} \right) \quad (2.49)$$

した場合について logit_i についての推定値 である。推定された logit^\wedge から元の死亡率を推定するためには、式 (2.42) を π について解いた逆ロジットの次式

$$\begin{aligned} \hat{\pi}_i &= \frac{\exp(\text{logit}_i^\wedge)}{1 + \exp(\text{logit}_i^\wedge)} \\ &= \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \end{aligned} \quad (2.50)$$

で推定することになる。逆ロジットは、死亡率が 0.0 から 1.0 の間でシグモイド曲線をあてはめる。この例では、160 歳強でほぼ 1.0 となることが計算されている。死亡者数が得られた範囲内に限定すれば、オフセット付きのポアソン回帰でもロジスティック回帰でも、ほとんど同じパラメータの推定値が得られるので、どちらが優れた方法とは言い難い。観測された死亡者数 y_i に対して統計モデルをあてはめるという観点からは、オフセット付きのポアソン回帰が望ましいように思われる。

表 2.25 分布を 2 項分布，リンク関数をロジットとした場合の結果

年齢層	死亡者数	母集団	死亡率	ロジット	推定値	逆ロジット
x	y	n	%	logit	logit^\wedge	%
30	1	17742	0.0056	-9.7836	-8.4985	0.0204
40	5	16059	0.0311	-8.0743	-7.4515	0.0580
50	25	10784	0.2318	-6.0646	-6.4045	0.1651
60	54	10706	0.5044	-5.2845	-5.3575	0.4691
80					-3.2635	3.6845
100					-1.1695	23.6945
120					0.9245	71.5958
140					3.0185	95.3403
160					5.1125	99.4015
		$\hat{\beta}_0 =$	-11.6395	$\hat{\beta}_1 =$	0.1047	

死亡率の上限を新たな変数としたロジスティック回帰

ロジスティック回帰では、死亡率が (0%~100%) の範囲のシグモイド曲線のあてはめを前提にしている。発想を変えて、死亡率の上限をデータから推定することも Excel のソルバーを用いると容易に推定することができる。

式 (2.51) は、下限を 0.0，上限を 1.0 とするロジスティック曲線であるが、下限も上限も変数として与えることができる。下限を L_{limit} ，上限を U_{limit} とするロジスティック曲線は、

$$\hat{\pi}_i = L_{limit} + \frac{(U_{limit} - L_{limit}) \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)} \quad (2.51)$$

である。一般化線形モデルとしてパラメータの推定はできないが、ソルバーならば推定可能である。表 2.26 に通常のロジスティック曲線のパラメータ推定値と上限付きのロジスティック曲線で、 $\hat{U}_{limit}=0.0074$ として推定されている。図 2.5 にそれぞれの曲線をプロットした結果を示す。対数尤度の比較からも通常のロジスティック曲線雄あてはめは支持されない。

表 2.26 上限を持つロジスティック曲線のあてはめ

					$\beta_0=$	-11.6398			$\beta_0=$	-10.1994	
					$\beta_1=$	0.1047	lnL		$\beta_1=$	0.1862	lnL
					$U_{limit}=$	-	-25.1465		$U_{limit}=$	0.0074	-19.0732
	年齢層	死亡者数	母集団	死亡率	上限が 1.0 のシグモイド				上限が U_{limit} のシグモイド		
i	x	y	人数 n	p	π^{\wedge}	B_p	$\ln L_i$		π^{\wedge}	B_p	$\ln L_i$
1	30	1	17,742	0.0001	0.0002	0.0972	-2.3314		0.0001	0.3557	-1.0336
2	35	5	16,554	0.0003	0.0003	0.1679	-1.7846		0.0002	0.1001	-2.3014
3	40	5	16,059	0.0003	0.0006	0.0523	-2.9502		0.0004	0.1245	-2.0838
4	45	12	13,083	0.0009	0.0010	0.1114	-2.1948		0.0010	0.1059	-2.2453
5	50	25	10,784	0.0023	0.0017	0.0221	-3.8127		0.0021	0.0738	-2.6062
6	55	38	9,645	0.0039	0.0028	0.0084	-4.7756		0.0038	0.0620	-2.7805
7	60	54	10,706	0.0050	0.0047	0.0475	-3.0464		0.0053	0.0496	-3.0045
8	65	65	9,933	0.0065	0.0079	0.0143	-4.2508		0.0064	0.0489	-3.0178
	70				0.0133				0.0070		
	80				0.0369				0.0073		
	100				0.2374				0.0074		
	200				0.9999				0.0074		

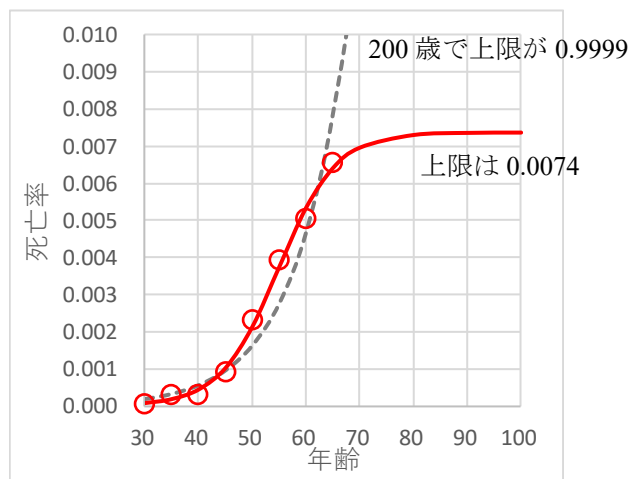


図 2.5 上限をパラメータとしたロジスティック曲線

第2章 文献索引

高橋(2015)	- 寿命試験データの統計解析	70
高橋(2017)	- 一般化線形モデルをExcelで極め活用するープロビット法・ロジット法・補2重対数法ー	70
高橋(2018)	- 正規分布を仮定した打ち切りデータを含む回帰分析入門	70

第2章 索引

あ	アドイン - ソルバー	69	最尤法 - 対数尤度	65
	アーミテジ (2001) - 冠動脈心疾患	88	- 反復重み付き回帰	68
	R言語 - optim() 最適化関数	68	指数関数 - 線形化	84
	1万人比 - オフセット	92	死亡者数 - 冠動脈心疾患	88
	位置パラメータ - ミューμ	66	死亡率 - 上限	93
	一般化線形モデル - JMP	74	JMP - 一般化線形モデル	74
	WolframAlpha - 数学ソフト	77	- 計算式エディタ	76
	- 偏微分	77	- 微分の機能	76
	打ち切りデータ - 高橋 (2018)	70	- 偏微分	75
	- ニュートン・ラフソン法	70	- 重ね合わせプロット	84
	Excel - SumProduct() 関数	75	- 対数リンク	87
	- ソルバー	68	- ポアソン回帰	82
	- 反復計算	86	- maximize() 最適化関数	68
	- 平滑化	67	種子数 - 久保 (2012)	84
	- Poisson.dist() 関数	63	種子の数 - 有害雑草	63
	SumProduct() 関数 - Excel	75	寿命試験データ - 高橋 (2015)	70
	(-H)-1 - 負の逆行列	75	- ワイブル回帰	70
	Hヘッセ - 2階の偏微分行列	70	上限 - 死亡率	93
	L - 尤度	64	- ロジスティック曲線	94
	ln L - 対数尤度	64	情報行列 - Fisherの情報量	72
	オフセット - 1万人比	92	- ヘッセ行列	72
	- 回帰式	90	初期値 - パラメータβ	80
	- 切片	90	人工データ - ドブソン (2008)	77
	- 対数リンク	88	- ドブソン (2008)	78
	optim() 関数 - 最適化	68	- ポアソン回帰	79
	optim() 最適化関数 - R言語	68	人口統計 - 母集団の人数	88
か	回帰式 - オフセット	90	推定値 - ポアソン回帰	65
	確率 - 尤度	64	数学ソフト - WolframAlpha	77
	確率P - 尤度L	64	スコアベクトルU - ヘッセ行列H	80
	確率関数 - 尤度関数	64	スネデカーら (1972) - 有害雑草	63
	下限・上限 - ロジスティック曲線	94	正規分布 - 打ち切りデータ	70
	重ね合わせプロット - JMP	84	正規方程式 - 偏微分	71
	冠動脈心疾患 - アーミテジ (2001)	88	切片 - オフセット	90
	- 死亡者数	88	切片のみ - ポアソン回帰	74
	逆ロジット - ロジット変換	93	線形化 - 指数関数	84
	共分散行列 - パラメータ	70	- リンク関数	84
	共変量 - 複数	83	ソルバー - アドイン	69
	久保 (2012) - 種子数	84	- Excel	68
	計算式エディタ - JMP	76	- 最大化	69
	恒等 - リンク関数	74	- 分析ツール	69
	恒等リンク - ポアソン回帰	78	た 対数リンク - オフセット	88
	誤差平方和 - Se	71	- JMP	87
さ	最大化 - ソルバー	69	- 偏微分	85
	- 対数尤度	63	- ポアソン回帰	84
	- 対数尤度	67	対数尤度 - ln L	64
	- 逐次的	63	- 最大化	63
	- ニュートン・ラフソン法	68	- 最尤法	65
	最適化 - optim() 関数	68	- 最大化	67
	最尤解 - 挟み撃ち法	66	対数尤度関数 - 2階の偏微分行列	68

- ニュートン・ラフソン法	68
- 偏微分	70
高橋 (2015) - 寿命試験データ	70
高橋 (2017) - ロジット法	70
高橋 (2018) - 打ち切りデータ	70
逐次的 - 最大化	63
ドブソン (2008) - 人工データ	77
- 人工データ	78
true - Poisson.dist() 関数	64
な 2階 - 偏微分行列	68
2階の偏微分行列 - Hヘッセ	70
- 対数尤度関数	68
- ヘッセ行列	70
2項分布 - ロジット	93
ニュートン・ラフソン法 - 打ち切りデータ	70
- 最大化	68
- 対数尤度関数	68
- 反復計算	72
- 反復過程	74
- 反復計算の実際	80
- ポアソン回帰	72
- ワイブル回帰	70
は 挟み撃ち法 - 最尤解	66
パラメータ - 共分散行列	70
- 分散	70
パラメータ β - 初期値	80
- 偏微分	79
パラメータ μ - 偏微分	71
反復過程 - ニュートン・ラフソン法	74
反復計算 - Excel	86
- ニュートン・ラフソン法	72
反復計算の実際 - ニュートン・ラフソン法	80
反復重み付き回帰 - 最尤法	68
微分の機能 - JMP	76
Fisherの情報量 - 情報行列	72
false - Poisson.dist() 関数	64
複数 - 共変量	83
複数の共変量 - ポアソン回帰	83
負の逆行列 - (-H)-1	75
分散 - パラメータ	70
分析ツール - ソルバー	69
平滑化 - Excel	67
ヘッセ行列 - 情報行列	72
- 2階の偏微分行列	70
ヘッセ行列H - スコアベクトルU	80
偏微分 - WolframAlpha	77
- JMP	75
- 対数尤度関数	70
- 対数リンク	85
- パラメータ μ	71
- パラメータ β	79
偏微分ベクトル - U	70
偏微分行列 - 2階	68
Poisson.dist() 関数 - Excel	63
- true	64
- false	64
ポアソン回帰 - 恒等リンク	78

- JMP	82
- 人工データ	79
- 推定値	65
- 対数リンク	84
- ニュートン・ラフソン法	72
- 複数の共変量	83
ポアソン確率 - 有害種子	68
母集団の人数 - 人口統計	88
ま maximize() 最適化関数 - JMP	68
μ - 位置パラメータ	66
や U - 偏微分ベクトル	70
有害雑草 - 種子の数	63
尤度 - L	64
- 確率	64
尤度L - 確率P	64
尤度関数 - 確率関数	64
有害雑草 - スネデガーら (1972)	63
有害種子 - ポアソン確率	68
ら リンク関数 - 恒等	74
- 線形化	84
- ロジット	93
ロジスティック回帰 - 一般線形モデル	93
ロジスティック曲線 - 下限・上限	94
- 上限	94
ロジット - 2項分布	93
- リンク関数	93
ロジット変換 - 逆ロジット	93
ロジット法 - 高橋 (2017)	70
わ ワイブル回帰 - 寿命試験データ	70
- ニュートン・ラフソン法	70

第2章 解析用ファイル一覧

	3 KB	第2章01_ポアソン確率	JMP Data Table
	45 KB	第2章01_種子数_尤度関数	Microsoft Excel ワークシート
	13 KB	第2章02_種子数_ソルバー	Microsoft Excel ワークシート
	18 KB	第2章03_種子数_ニュートン	Microsoft Excel ワークシート
	3 KB	第2章03_種子数_偏微分	JMP Data Table
	18 KB	第2章04_人工データ	Microsoft Excel ワークシート
	4 KB	第2章04_人工データ_偏微分式	JMP Data Table
	10 KB	第2章05_久保_種子_グラフ化	JMP Data Table
	9 KB	第2章05_久保_種子_偏微分式	JMP Data Table
	25 KB	第2章05_久保_種子C群	Microsoft Excel ワークシート
	3 KB	第2章06_冠動脈疾患_2項分布	JMP Data Table
	43 KB	第2章06_冠動脈疾患_オフセット	Microsoft Excel ワークシート
	3 KB	第2章06_冠動脈疾患_偏微分式	JMP Data Table
	3 KB	第2章06_冠動脈疾患グラフ	JMP Data Table

非売品, 無断複製を禁ずる

第 9 回 続高橋セミナー

最尤法によるポアソン回帰分析入門<<第 2 章>>

第 2 章 ニュートン・ラフソン法によるポアソン回帰

BioStat 研究所(株)

〒105-0014 東京都 港区 芝 1-12-3 の1005

2020 年 4 月 7 日 高橋 行雄