

# カブトガニのサテライト数の観察データに関する探索的ポアソン回帰

高橋 行雄  
BioStat 研究所(株)

Exploratory Poisson regression on observation data of the number of satellites in horseshoe crabs

Yukio Takahashi  
BioStat Research Co.,Ltd.

**要旨：** 観察データとして得られたカウント・データに対し、ポアソン回帰による探索的な解析を試みようとしたときに(恒等リンク or 対数リンク)、(オフセットの有り or 無し)、(過分散の調整 or ガンマ・ポアソン分布 or ゼロ過剰ポアソン分布)などの選択が迫られる。同じ回帰分析なのに、なれ親しんできた正規分布を仮定した回帰分析とは、全く異次元の世界のごとくである。そこで、成書で取り上げられている「カブトガニのサテライト数の観察データ(付録A)」に対する探索的な解析方法を示すことにより、ポアソン回帰について理解の向上を図りたい。このデータは、173 匹のカブトガニについて説明変数として順序尺度データ(甲羅の色、後体部の棘の状態)の 2 変数、連続尺度(甲羅の幅、体重)の 2 変数、応答変数としてサテライト数が含まれている。全体としてはゼロ過剰ガンマ・ポアソン分布のあてはめがよいが、探索的な解析での分布としては疑問が残る。また、2 つの順序尺度間の交互作用解析は、セル度数の分布が均一ではないという問題がある。さらに、2 つの連続変数に高い相関があり、選択の問題もある。これらの探索的な解析に伴う種々の問題に対し、JMPには「一般化線形モデル」にプロファイル機能が備わり、交互作用の検討などに威力を発揮する。また、「グラフ・ビルダー」に散布図行列中に回帰直線と 95%信頼区間を上書きできる機能があり、これまで十分とは思えなかったサテライト数におよぼす探索的解析を行なったので結果を示す。

**キーワード：** ポアソン回帰、最尤法、過分散、一般化線形モデル、JMP、SAS/GENMOD

## 1. はじめに

SAS の GENMOD プロシジャにより一般化線形モデルが使えるようになった時に、最初に注目したのは、従来からある LOGISTIC プロシジャとの使い分けであった。その詳細は、高橋(2002)に示したが、ポアソン回帰については、全く扱っていなかった。その後、高橋(2004)で、細胞毒性データについて、GENMOD プロシジャを用いたポアソン回帰による勾配比検定についての検討結果を示した。

久保(2012)は、「何でもかんでも正規分布と考えるのはおかしいだろう」というコンセプトで正規分布ではなくポアソン分布を全面的に取り上げて論じている。さらに、植物の種子数を主体した「一般化線形モデル(GLM)ーポアソン回帰ー」を展開し、これまでの正規分布を前提とした統計解析とは異なる切り口を提示した。これに刺激を受け、高橋(2004)をベースにし、「最尤法によるポアソン回帰入門」と題して大幅な改訂作業を行っている。

その一環として、アグレスティ著、渡邊・菅波・吉田ら訳(2003)に示されている雌のカブトガニに連結する雄のサテライト数(Satellite 数)の例を取り上げる。このデータには、付表 A に示すように、173 匹のカブトガ

ニについて、説明変数として順序尺度データ(甲羅の色、後体部の棘の状態)の 2 変数、連続尺度(甲羅の幅、体重)の 2 変数、応答変数としてサテライト数が含まれている。文献では、甲羅の幅をX軸、サテライト数をY軸とした散布図と共に、対数リンクによるポアソン回帰の結果が示されている。しかし、その後の解析では甲羅の幅を 8 区分とし区分内のカブトガニの数とサテライト数の合計を算出し、カブトガニの数をオフセットとした解析を主体にしている。探索的な解析では、サテライト数が(0, 1 以上)の 2 値データとして、ロジスティック化回帰を主体にした解析方法が提示されている。

## 2. ポアソン分布のあてはめ

表 1 に示すように、JMP の「一変量の分布」によりサテライト数の平均は 2.9191、分散は 9.9120 であり、その比は 3.40 と過分散になっている。ポアソン分布をあてはめ、棒グラフ上に上書きした結果を見ても、誤差分布にポアソン分布を仮定することは絶望的とも思われる。もちろん、適合度の検定でも  $\chi^2 = 584.0436$ ,  $p < 0.0001$  でポアソン分布があてはまるとは言えない。このような全データで過分散となる場合では、何らかの条件によりサテライト数の平均が大きく異なる部分集団の集まりが複数存在する可能性も考えられる。

表 1 サテライト数へのポアソン分布のあてはめ



## 3. 過分散を調整したポアソン回帰

甲羅の幅を説明変数とし、サテライト数を応答変数とした場合に、次式の対数リンク(両辺に対数を取った時に線形となる)によるポアソン回帰

$$Satellite_i = \exp(\beta_0 + \beta_1 \cdot width_i) + \varepsilon_i, \quad \varepsilon_i \sim Poisson$$

の結果を表 2 に示す。Pearson の適合度のカイ 2 乗値は 544.1570 と自由度の 171 に対して 3.1822 倍と過

表 2 甲羅の幅についての対数リンクの過分散調整なしのポアソン回帰

適合度統計量		カイ2乗	自由度	p値(Prob>ChiSq)
Pearson		544.1570	171	<.0001*
デビアン		567.8786	171	<.0001*

パラメータ推定値				
項	推定値	標準誤差	尤度比カイ2乗	p値
切片	-3.3048	0.5422	36.8670	<.0001*
甲羅の幅	0.1640	0.0200	64.9131	<.0001*

分散となっている。図 1 に示すように、ポアソン回帰から得られた尤度比カイ 2 乗値を自由度で除した過分散パラメータを  $\phi = 3.1822$  とし、得られた共分散行列に  $\phi$  倍して標準誤差を調整する方法が知られていて、JMP のポアソン回帰でもサポートされている。

手法: 一般化線形モデル

分布: Poisson

リンク関数 対数

☒ 過分散に基づく検定と信頼区間
 ☐ Firth バイアス調整推定値

図 1 過分散の調整法オプション

過分散を調整したポアソン回帰の結果を

表 3 に示す。表 2 に示した甲羅の幅の標準誤差は、 $SE = 0.0200$  であったので、調整後の  $SE'$  は、

$$SE' = \sqrt{\phi SE^2} = \sqrt{3.1822 \times 0.0200^2} = 0.0356$$

と大きくなり、尤度比カイ 2 乗値は、64.9131 から 20.3988 と激減する。

過分散の係数を用いた方法は、過分散となるカウント・データに対する万能の方法とも思われるかもしれないが、表 1 に示したヒストグラムに重ね書きしたポアソン分布から、このデータにポアソン分布を仮定することは全くできない。もちろん甲羅の幅に対するポアソン回帰で過分散が解消するのであれば嬉しいのであるが、実際にどのような分布になるのか示すことができない。単に  $SE$  を割り増ししているだけである。

表 3 過分散調整済みのポアソン回帰

項	推定値	標準誤差	尤度比カイ2乗	p値
切片	-3.3048	0.9673	11.5854	0.0007*
甲羅の幅	0.1640	0.0356	20.3988	<.0001*

#### 4. ポアソン回帰の個別データの 95%信頼区間

ポアソン回帰を行っても過分散が解消していないことを視覚化するために散布図に個別データの 95%信頼区間(予測区間)を重ね書きしてみると、図 2 左に示すように上側に多数の点のはみ出ているのでポアソン回帰(指数曲線)のあてはめには無理があることを実感できる。図 2 右に示すように予測値に対する Pearson 残差をプロットすることにより、Pearson 残差が 3 以上の飛び離れデータが多数存在することからも、ポアソン分布を誤差分布とする回帰分析について否定的な結果となっている。

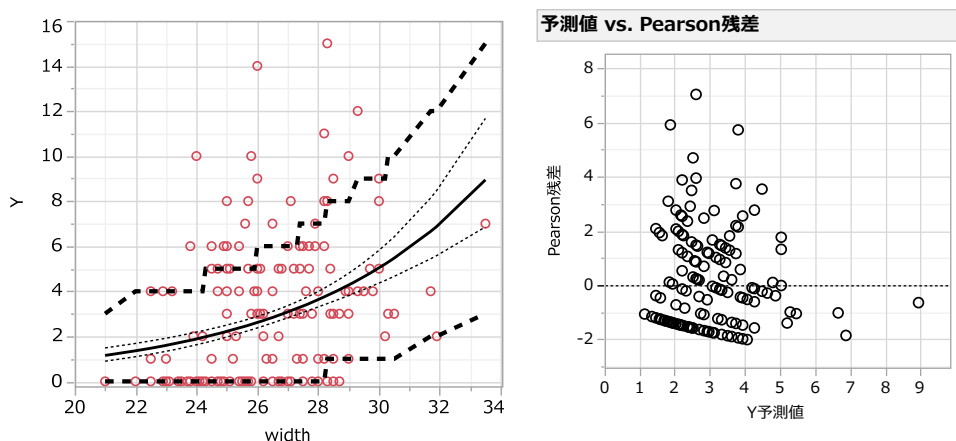


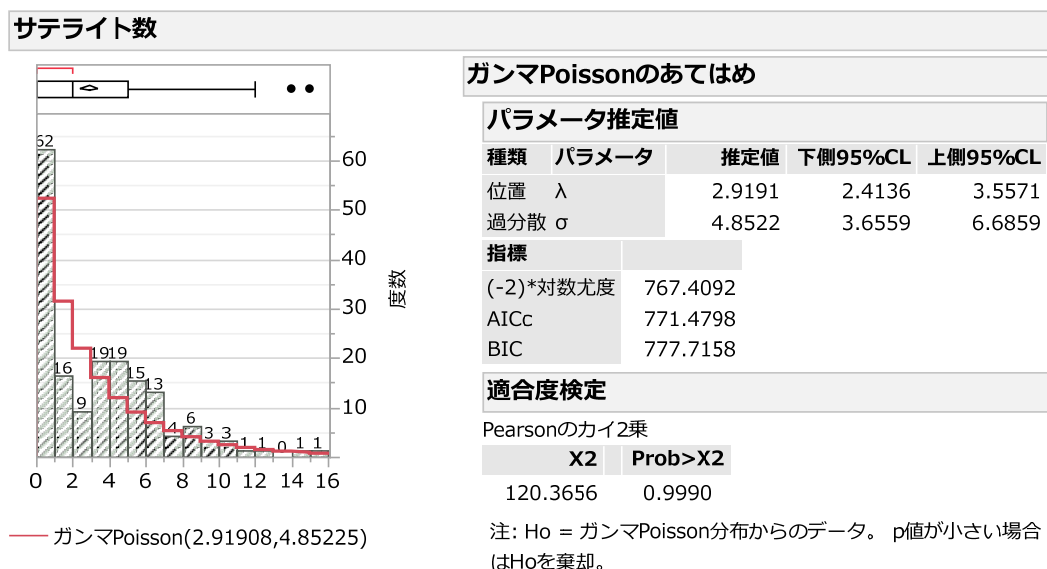
図 2 ポアソン回帰に対する 95%信頼区間および予測値に対する Pearson 残差

他の変数を加えてポアソン回帰を行っても過分散が解消されないのであれば、ポアソン回帰を行う前提がないことになる。主な原因は、173 個体に対してサテライト数ゼロが 62 匹と全体の 35.8%なので、サテライト数を (0, 1) 反応とする解析が望ましいとも考えられる。また、サテライト数が 3 と 4 あたりに分布の山があることから、あるいは、3 区分程度の順序データとする解析を行うことが望ましいかも知れない。

## 5. ガンマ・ポアソン分布のあてはめ

JMP には、過分散を考慮した負の二項分布から導出されたガンマ・ポアソン分布をあてはめる機能があるので、表 4 に結果を示す。結果は、位置  $\lambda = 2.9191$ 、過分散  $\sigma = 4.8522$  となる。表 1 に示したポアソン分布のあてはめでは、サテライト数がゼロの場合について大きな乖離があったが、過分散を考慮したガンマ・ポアソン分布では、まずまずのあてはめが行われているように思われる。

表 4 サテライト数へのガンマ・ポアソン分布のあてはめ



## 6. 層別解析

付録 A に示したデータには、説明変数として順序尺度(甲羅の色, 後体部の棘の状態)の 2 変数があるので JMP の「二変量の関係」で作成したサテライト数に対する層別分布を図 3 に示す。

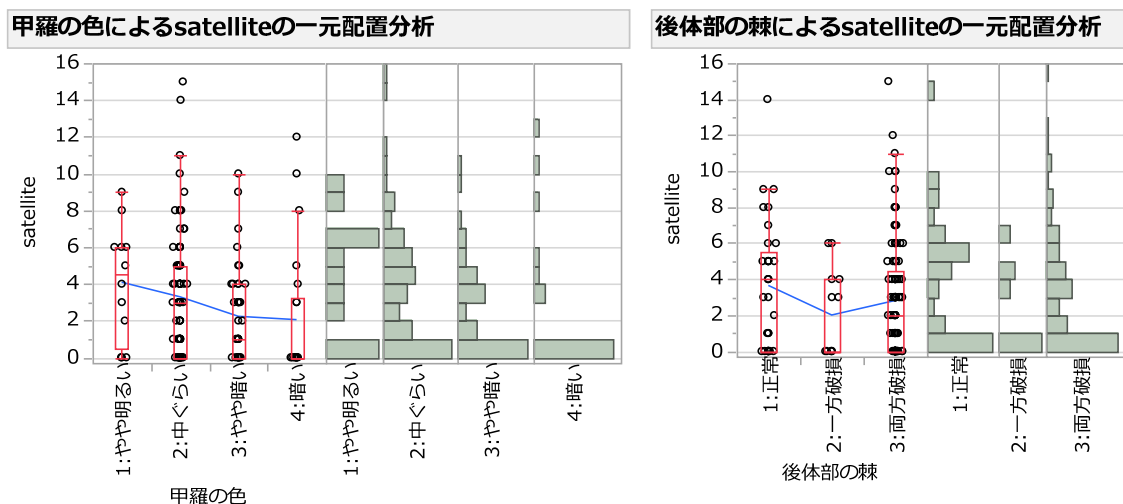


図 3 甲羅の色および後体部の棘の状態とサテライト数の関連

雌の甲羅の色については、暗くなるに従いゼロ・カウントが増えサテライト数の平均値が減少傾向であることが読み取れる。雌の後体部の棘の状態については、正常の場合には、サテライト数の 5 匹に山があり、雄が連結する割合が多いようであるが、サテライト数の平均値は同程度である。

表 5 に甲羅の色と後体部の棘の状態を組み合わせた場合のサテライト数 N、サテライト数の平均と分散を示す。甲羅の色が暗くなるにつれて後体部の棘は、正常から破損へ移行するが、ある程度のサテライト数がある場合の分散/平均の比は、2 以上あり過分散が解消する様子はない。

表 5 甲羅の色別 後体部の棘別 のサテライト数の分散/平均の比

甲羅の色	棘の状態	N	平均	分散	分散/平均
1:やや明るい	1:正常	9	4.44	10.53	2.37
	2:一方破損	2	4.50	4.50	1.00
	3:両方破損	1	0.00	-	-
2:中ぐらい	1:正常	24	3.29	12.13	3.68
	2:一方破損	8	1.75	6.21	3.55
	3:両方破損	63	3.49	10.03	2.87
3:やや暗い	1:正常	3	5.33	10.33	1.94
	2:一方破損	4	1.75	4.25	2.43
	3:両方破損	37	2.03	6.25	3.08
4:暗い	1:正常	1	0.00	-	-
	2:一方破損	1	0.00	-	-
	3:両方破損	20	2.25	13.99	6.22
	全体	173	2.92	9.91	3.40

## 7. 甲羅の幅か体重か

過分散を承知で、対数リンクによる 2 変数のポアソン回帰を行い、幅か体重か、どちらがサテライト数との関連が高いか検討する。表 6 に示すように、甲羅の幅の推定値は、0.0461、体重の推定値は、0.4470 であり、尤度比検定の結果は、体重のみが有意な差であった。

表 6 対数リンクによるポアソン 2 変量回帰

項	推定値	標準誤差	尤度比カイ2乗	p値
切片	-1.2952	0.8989	2.0691	0.1503
甲羅の幅	0.0461	0.0467	0.9658	0.3257
体重	0.4470	0.1586	7.9780	0.0047

図 4 は、JMP による対数リンクでの 2 変量ポアソン回帰に引き続き「予測プロファイル」の機能を用い、体重を(2, 3, 4 kg)と変化させた場合の甲羅の幅がサテライト数に及ぼす影響を図示したものである。甲羅の幅は体重の増加に伴いサテライト数も増加しているが、95%信頼区間の表示から、傾きがマイナスになる可能性があることが読み取れ、このことが表 6 の p 値が大きいことに対応する。

さて、甲羅の色が暗くなる日本語つれて棘の破損が多くなり、サテライト数が減ることを表 5 で示した。では、甲羅の色と体重を組み合わせた場合に、何らかの関連が見出されるのであろうか。このような関連を、ポアソン回帰で見い出すためには、甲羅の色について何らかの数値を与えてデザイン行列化し、体重との交互作用を含めたポアソン回帰を行う必要がある。JMP の一般化線形モデルでは、名義尺度に対しては対比型のデザイン行列を自動生成するので、「モデル効果の構成」で(甲羅の色, 体重, 甲羅の色\*体重)を設定すればよい。

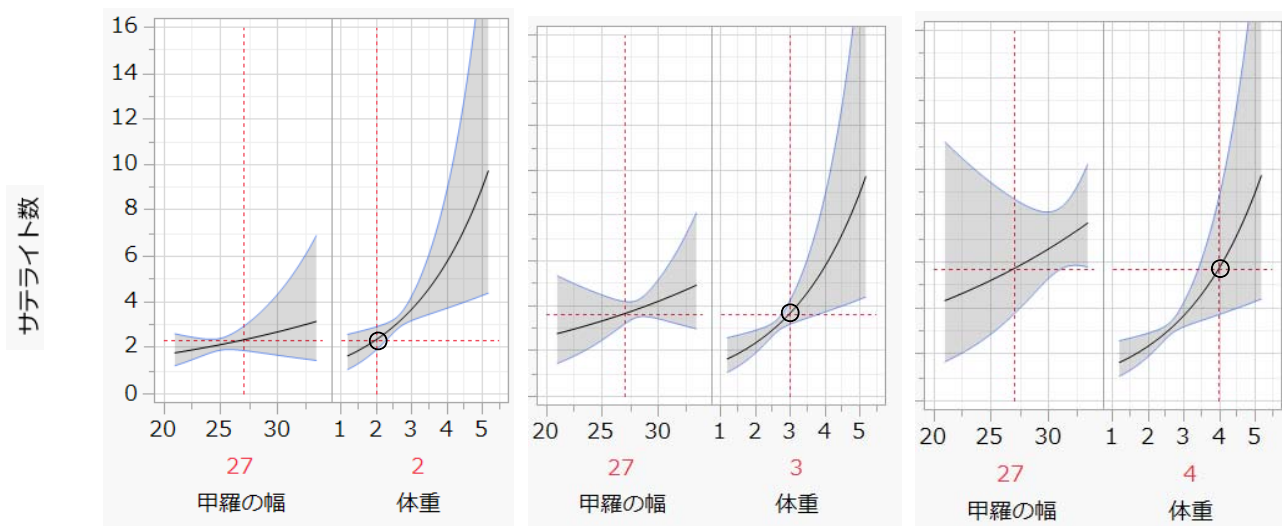


図4 体重を変化させた場合の甲羅の幅とサテライト数との関連

パラメータの推定結果を表7に示すが、このままでは、結果の解釈は困難を極めるので、「予測プロファイル」の機能を用いて図5に示すように甲羅の色ごとの体重の増加によるサテライト数との関連を概観する。

表7 甲羅の色と体重の交互作用を含めた対数リンクでのポアソン重回帰

項	推定値	標準誤差	尤度比カイ2乗	p値
切片	-0.2778	0.3450	0.6530	0.4191
甲羅の色[1:やや明るい]	2.2221	0.7978	7.5086	0.0061*
甲羅の色[2:中ぐらい]	0.2010	0.3797	0.2812	0.5959
甲羅の色[3:やや暗い]	-1.1855	0.4865	6.0352	0.0140*
体重	0.5463	0.1344	16.0804	<.0001*
甲羅の色[1:やや明るい]*体重	-0.7518	0.3050	6.1530	0.0131*
甲羅の色[2:中ぐらい]*体重	-0.0646	0.1456	0.1967	0.6574
甲羅の色[3:やや暗い]*体重	0.3820	0.1870	4.2010	0.0404*

予測プロファイルから、甲羅の色が「やや明るい」場合は、体重とサテライト数の関連は、マイナスの傾きも起こりえる信頼区間となっており、関連はみいだせないことが他の色と明らかに異なる。「中ぐらい」以上では、体重が増えればサテライト数も増大する。後体部の棘の状態は、甲羅の色によって破損が進行するこ

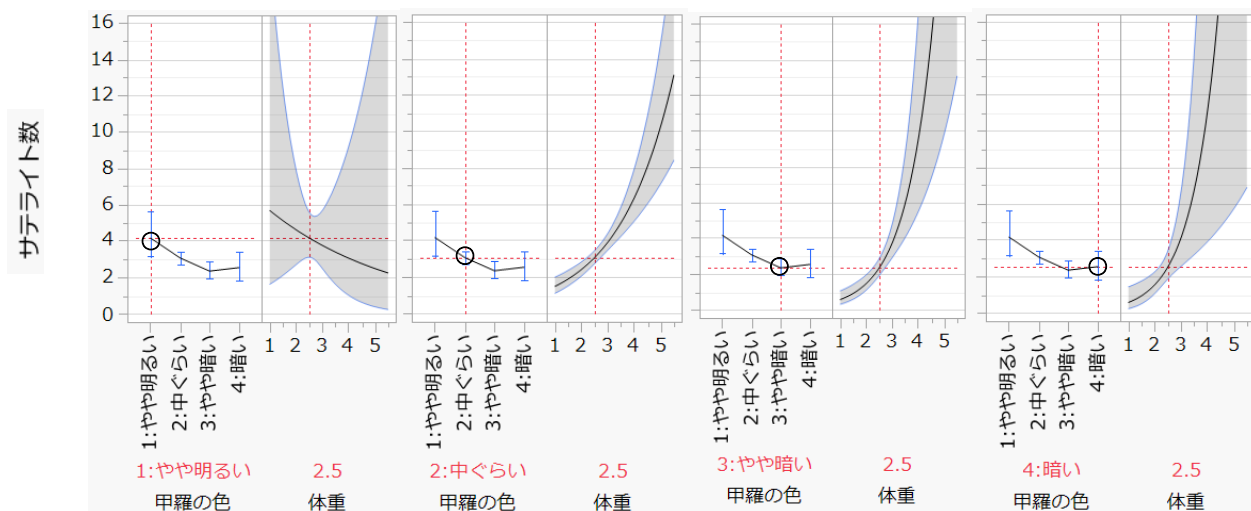


図5 甲羅の色を変化させた場合の体重とサテライト数の関連



とを表 5 で明らかにした。甲羅の色が「中ぐらい」の場合には、後部の棘が「正常」と「両方破損」に分かれているので、サテライト数との関連を甲羅の色が「中ぐらい」に限定して関連を調べた結果を図 6 に示す。

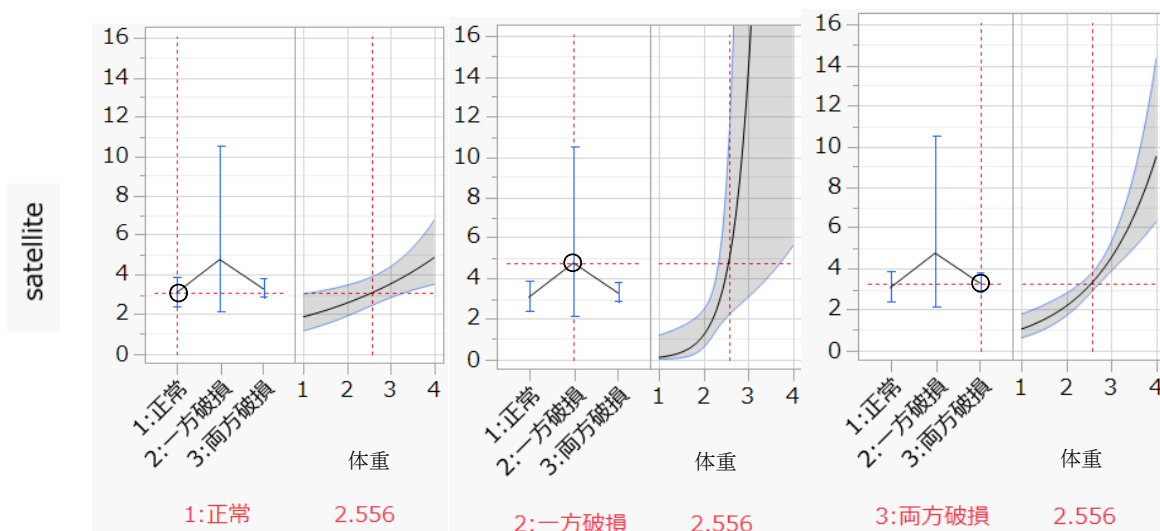


図 6 甲羅の色が「中ぐらい」での後部の棘の状態別の体重とサテライト数の関連

図 6 から甲羅の色が「中ぐらい」で後体部の棘が「正常」の場合に体重が増えればサテライト数も微増する。「一方破損」および「両方破損」では、体重が増えた場合にサテライト数が急増する。表 5 から、甲羅の色が「やや明るい」場合には、後体部の棘は 12 匹中 9 匹が「正常」で、図 5 から体重が増えてもサテライト数は増えない。甲羅の色が「中ぐらい」に変化すると、体重が増加するとサテライト数も大幅に増える。更に色が「やや暗い、暗い」場合には、更に体重が増えるにつれて、サテライト数が増えるとも言えるが、体重が小さい場合には、サテライト数が減少することが読み取れる。

## 8. グラフ・ビルダーによる散布図行列における回帰の 95%信頼区間

交互作用が疑われるような探索的な解析を行うためには、各種のグラフ表示が欠かせない。これまでも JMP の多彩なグラフ表示を活用し、カブトカニの各種の変数とサテライト数の関連を浮き彫りにしてきたが、満足できるものではなかった。全体を俯瞰できるように結果を 1 枚のグラフで表わすことは、可能なのだろうか。JMP の新しい作図機能である「グラフ・ビルダー」を用いた結果を図 7 に示す。

この図から、これまでの探索的な解析の結果がより鮮明に浮彫される。サテライト数は、甲羅の色が暗くなるにつれて後体部の棘の破損が進み、それに伴い、体重の軽い雌ほど連結する雄のサテライト数が減少することが読み取れる。甲羅の色が暗くなり、後体部の棘の状態が悪くなる加齢現象により、体重の軽い雌ほど連結する雄のサテライト数が減少すると解される。そのため、ゼロ・カウントが多い過分散となったと推測される。

## 9. 考察

一般化線形モデルで定式化されているポアソン回帰については、ドブソン著、田中・森川・山中ら訳 (2008) に丁寧な解説があり必読の書である。反復重み付き回帰による最尤法によるポアソン回帰について、人工データを用いた丁寧な導入があり、これに基づき高橋 (2019) で Excel を用いたポアソン回帰による勾配検定を行うことができた。

通常の回帰分析に対応したポアソン回帰は、恒等リンクとした場合であるが、各種の応用例で散見する

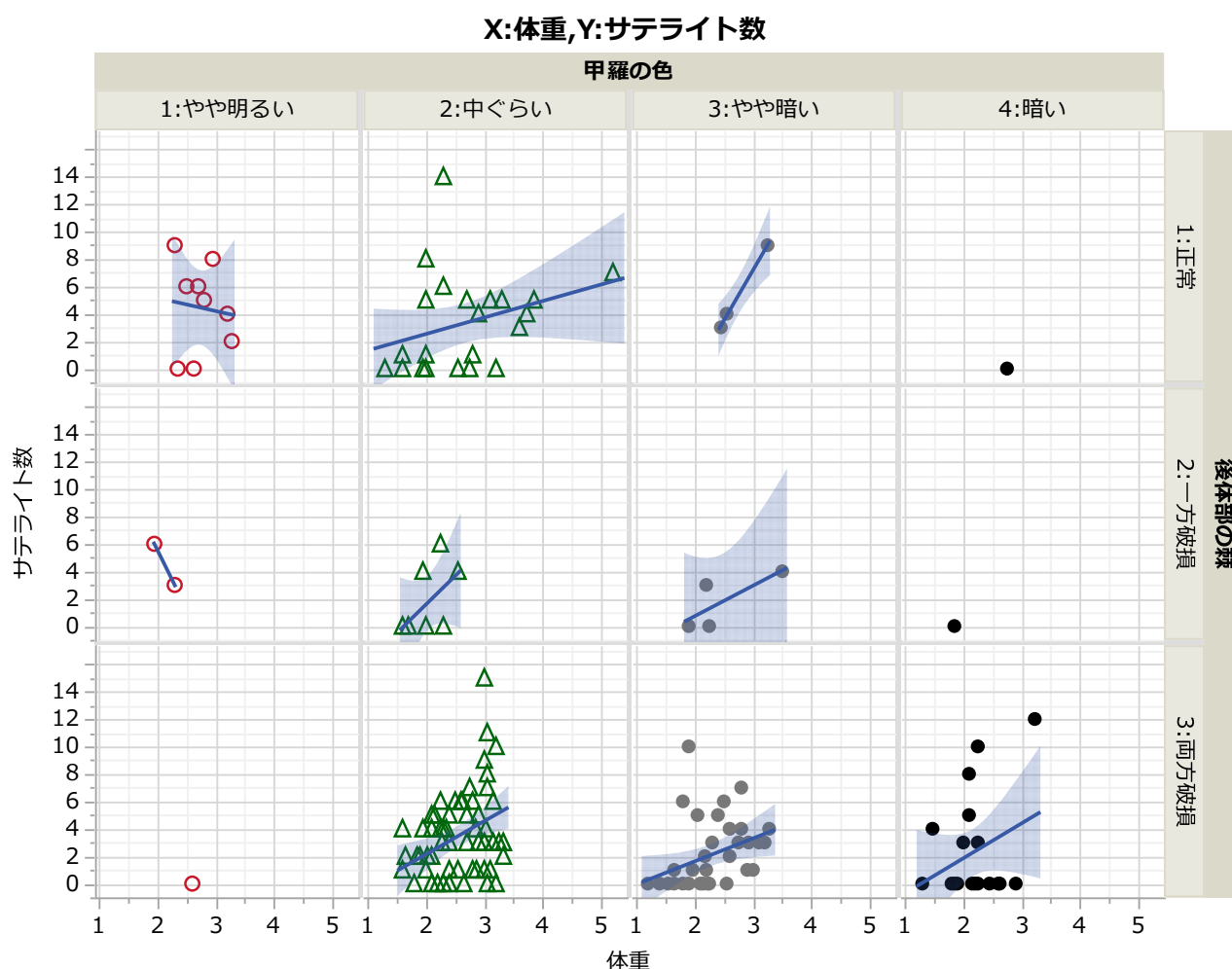


図7 甲羅の色・棘の状態による層別散布図での回帰の95%信頼区間の表示

のは、対数リンクでオフセットがあり、あるいは過分散を考慮する場合など多彩である。JMP の一般化線形モデル、SAS の GENMOD プロシジャでポアソン回帰ができるようになっているが、適当な参考文献は見当たらない。ドブソンの訳本でも、ポアソン回帰の例示は、対数リンクでオフセットがあり、2 乗項もあり、さらに交互作用も含めた事例が示されている。もちろん統計ソフトを使う前提で、追試も容易にできるのだが、その結果の解釈は難解である。

アグレスティの訳本で例示されているカブトガニのサテライト数の事例は、通常の回帰分析と対比しやすいので探索的なポアソン回帰の例示として取り上げた。全データに対するサテライト数の分布について、ゼロ過剰ポアソン分布よりも、さらにゼロ過剰ガンマ・ポアソン分布のあてはめが良好であったが、それらの分布を用いた回帰分析には難点がある。これは、図 2 にも示したように、甲羅の幅が大きい時にはサテライト数のゼロが存在しなくなるので、ゼロ過剰ガンマ・ポアソン分布を仮定して回帰分析を行うと、体重が重い場合にも過剰なゼロが存在を仮定することになり、現実のデータとの乖離を無視できなくなるためである。

対数リンクによるポアソン回帰は、元データには指数曲線のあてはめ、両辺に对数を取るモデルであり、ゼロ・データに対しては対数変換が行われないように調整する仕組みになっている。この仕組みは、一般化線形モデルで分布を正規とし、対数リンクとした場合でも適用され、ゼロを含むようなデータに対し指数曲線をあてはめることが可能となる。なお、ポアソン回帰を行っても過分散が解消されないような場合に、正規分布を仮定し、対数リンクによる指数曲線をあてはめる場合にも、ゼロ・データに対する調整が行われる。



探索的ポアソン回帰は、表 1 でも示したように過剰なゼロが、どのような状況で発生するかを念頭にし、「甲羅の色」、「後体部の棘」とサテライト数の関係から、甲羅の色が暗くなるにつれゼロ・カウントが増加するが、後体部の棘については、関連が見いだされなかった。さらに、甲羅の色と後体部の棘を組み合わせても過分散は解消しなかった。

甲羅の幅と体重の 2 変数間には 0.89 と高い相関があり、2 変数のポアソン回帰に引き続き、図 4 に示したように体重を段階的に変化させた場合の甲羅の幅の推定曲線と 95%信頼区間のプロファイルから、甲羅の幅をポアソン回帰の説明変数に加える必要がないことが、視覚的に見いだされた。もちろん、2 変数のポアソン回帰の尤度比検定で、甲羅の幅の  $p$  値は 0.3257 と有意ではないことから推測されることではあるが、JMP のプロファイル機能は、視覚的に変数相互の関連を見出し、より具体的な相互関係の理解するために有益である。

このプロファイル機能により、図 5 に示したように 4 水準の甲羅の色と体重の 2 変数に交互作用を加えたポアソン回帰で、甲羅の色が「やや明るい」場合に、体重が増えてもサテライト数が増えないことが図示され、甲羅の色が「中ぐらい、やや暗い、暗い」場合とは、全く異なるプロファイルであることが明示された。他方、図 6 に示すように後体部の棘と体重の関連には、交互作用を示唆するような兆候は見いだせなかった。

甲羅の色と後体部の棘に体重、さらにそれらの交互作用を含めたポアソン回帰は、観察データなので、データが不均一であり、解を得ることができなかった。これらの変数とサテライト数の関連を見出すためには、図 7 に示すように JMP のグラフ・ビルダーが役に立つ。最初に体重とサテライト数の散布図を描き、回帰直線と 95%信頼区間を上書きする。ここまでならば、JMP の伝統的な二変量の関係での対応と同じであるが、これに 4 水準の甲羅の色、3 水準の後体部を組み合わせた  $4 \times 3$  の場合についてタイル状に体重とサテライト数の回帰直線と 95%信頼区間を並べて表示できた。グラフ・ビルダーで対数リンクのポアソン回帰が実施できれば申し分ないのであるが、残念ながら現在のバージョン 14 では対応していない。

伝統的な回帰分析であっても、名義尺度の水準ごとの散布図行列上に回帰直線の 95%信頼区間が表示されるだけでも、結果を総合的に俯瞰するために有益である。これに類似する機能が S プラスにあり、以前は愛用していたのであるが、JMP グラフ・ビルダーは、S プラスの機能を大幅に凌駕する探索的な統計解析を支援するツールとして優れている。

## 参 考 文 献

- 1) 高橋行雄(2002), GENMODプロシジャによる計数データの解析, SAS ユーザー総会論文集:193-202.
- 2) 高橋行雄(2004), ポアソン回帰分析入門ー細胞数をカウントしたデータの解析ー, <https://www.yukms.com/biostat/takahasi/rec/017.htm> 2019 年 7 月 19 日アクセス.
- 3) 久保拓弥(2012), データ解析のための統計モデリング入門 一般化線形モデル・階層ベイズモデル・MCMC, 岩波書店:39-65.
- 4) アグレスティ著, 渡邊裕之・菅波秀樹・吉田光弘ら訳(2003), カテゴリカルデータ解析入門サイエンティスト社:110-127, 168-179.
- 5) ドブソン著, 田中豊・森川義彦・山中竹春・富田誠 訳(2008), 一般化線形モデル入門, 原著 第 2 版, 共立出版:67-80, 186-189.
- 6) 高橋行雄(2019), ポアソン回帰を用いた勾配比検定, 2019 年度日本計量生物学会講予稿集;65-70.

付録 A 雌のカブトガニに連結する雄のサテライト数

col or	spi ne	width	weight	sat ell	col or	spi ne	width	weight	sat ell	col or	spi ne	width	weight	sat ell	col or	spi ne	width	weight	sat ell
2	3	28.3	3.050	8	3	1	28.5	3.250	9	4	3	23.5	1.900	0	2	1	28.0	2.900	4
3	3	22.5	1.550	0	3	3	28.9	2.800	4	2	2	24.0	1.700	0	4	3	25.8	2.250	10
1	1	26.0	2.300	9	2	3	28.2	2.600	6	2	1	29.7	3.850	5	2	3	27.9	3.050	7
3	3	24.8	2.100	0	2	3	25.0	2.100	4	2	1	26.8	2.550	0	2	3	24.9	2.200	0
3	3	26.0	2.600	4	2	3	28.5	3.000	3	4	3	26.7	2.450	0	2	1	28.4	3.100	5
2	3	23.8	2.100	0	2	1	30.3	3.600	3	2	1	28.7	3.200	0	3	3	27.2	2.400	5
1	1	26.5	2.350	0	4	3	24.7	2.100	5	3	3	23.1	1.550	0	2	2	25.0	2.250	6
3	2	24.7	1.900	0	2	3	27.7	2.900	5	2	1	29.0	2.800	1	2	3	27.5	2.625	6
2	1	23.7	1.950	0	1	1	27.4	2.700	6	3	3	25.5	2.250	0	2	1	33.5	5.200	7
3	3	25.6	2.150	0	2	3	22.9	1.600	4	3	3	26.5	1.967	1	2	3	30.5	3.325	3
3	3	24.3	2.150	0	2	1	25.7	2.000	5	3	3	24.5	2.200	1	3	3	29.0	2.925	3
2	3	25.8	2.650	0	2	3	28.3	3.000	15	3	3	28.5	3.000	1	2	1	24.3	2.000	0
2	3	28.2	3.050	11	2	3	27.2	2.700	3	2	3	28.2	2.867	1	2	3	25.8	2.400	0
4	2	21.0	1.850	0	3	3	26.2	2.300	3	2	3	24.5	1.600	1	4	3	25.0	2.100	8
2	1	26.0	2.300	14	2	1	27.8	2.750	0	2	3	27.5	2.550	1	2	1	31.7	3.725	4
1	1	27.1	2.950	8	4	3	25.5	2.250	0	2	2	24.7	2.550	4	2	3	29.5	3.025	4
2	3	25.2	2.000	1	3	3	27.1	2.550	0	2	1	25.2	2.000	1	3	3	24.0	1.900	10
2	3	29.0	3.000	1	3	3	24.5	2.050	5	3	3	27.3	2.900	1	2	3	30.0	3.000	9
4	3	24.7	2.200	0	3	1	27.0	2.450	3	2	3	26.3	2.400	1	2	3	27.6	2.850	4
2	3	27.4	2.700	5	2	3	26.0	2.150	5	2	3	29.0	3.100	1	2	3	26.2	2.300	0
2	2	23.2	1.950	4	2	3	28.0	2.800	1	2	3	25.3	1.900	2	2	1	23.1	2.000	0
1	2	25.0	2.300	3	2	3	30.0	3.050	8	2	3	26.5	2.300	4	2	1	22.9	1.600	0
2	1	22.5	1.600	1	2	3	29.0	3.200	10	2	3	27.8	3.250	3	4	3	24.5	1.900	0
3	3	26.7	2.600	2	2	3	26.2	2.400	0	2	3	27.0	2.500	6	2	3	24.7	1.950	4
4	3	25.8	2.000	3	2	1	26.5	1.300	0	3	3	25.7	2.100	0	2	3	28.3	3.200	0
4	3	26.2	1.300	0	2	3	26.2	2.400	3	2	3	25.0	2.100	2	2	3	23.9	1.850	2
2	3	28.7	3.150	3	3	3	25.6	2.800	7	2	3	31.9	3.325	2	3	3	23.8	1.800	0
2	1	26.8	2.700	5	3	3	23.0	1.650	1	4	3	23.7	1.800	0	3	2	29.8	3.500	4
4	3	27.5	2.600	0	3	3	23.0	1.800	0	4	3	29.3	3.225	12	2	3	26.5	2.350	4
2	3	24.9	2.100	0	2	3	25.4	2.250	6	3	3	22.0	1.400	0	2	3	26.0	2.275	3
1	1	29.3	3.200	4	3	3	24.2	1.900	0	2	3	25.0	2.400	5	2	3	28.2	3.050	8
1	3	25.8	2.600	0	2	2	22.9	1.600	0	3	3	27.0	2.500	6	4	3	25.7	2.150	0
2	2	25.7	2.000	0	3	2	26.0	2.200	3	3	3	23.8	1.800	6	2	3	26.5	2.750	7
2	1	25.7	2.000	8	2	3	25.4	2.250	4	1	1	30.2	3.275	2	2	3	25.8	2.200	0
2	1	26.7	2.700	5	3	3	25.7	1.200	0	3	3	26.2	2.225	0	3	3	24.1	1.800	0
4	3	23.7	1.850	0	2	3	25.1	2.100	5	2	3	24.2	1.650	2	3	3	26.2	2.175	2
2	3	26.8	2.650	0	3	2	24.5	2.250	0	2	3	27.4	2.900	3	3	3	26.1	2.750	3
2	3	27.5	3.150	6	4	3	27.5	2.900	0	2	2	25.4	2.300	0	3	3	29.0	3.275	4
4	3	23.4	1.900	0	3	3	23.1	1.650	0	3	3	28.4	3.200	3	1	1	28.0	2.625	0
2	3	27.9	2.800	6	3	1	25.9	2.550	4	4	3	22.5	1.475	4	4	3	27.0	2.625	0
3	3	27.5	3.100	3	2	3	25.8	2.300	0	2	3	26.2	2.025	2	2	2	24.5	2.000	0
1	1	26.1	2.800	5	4	3	27.0	2.250	3	2	1	24.9	2.300	6					
1	1	27.7	2.500	6	2	3	28.5	3.050	0	1	2	24.5	1.950	6					
2	1	30.0	3.300	5	4	1	25.5	2.750	0	2	3	25.1	1.800	0					

注釈: color=色(1=やや明るい, 2=中くらい, 3=やや暗い, 4=暗い);

Spine=後体部の棘の状態(1=いずれも正常, 2=一方が摩耗または破損している, 3=いずれも摩耗または破損している);

width=甲羅の幅(cm); weight=重さ(kg); satell=サテライト数.

出典: <http://lib.stat.cmu.edu/datasets/agresti>. 2019年7月24日 アクセス.