

## 医学情報処理演習第12回「一般化線型モデル」課題回答例

必要な計算をするプログラムは以下の通りである。変数選択は必要ない。

```
it12-ans-2006.R  
  
NagelkerkeR2 <- function(rr) {  
  print(n <- nrow(rr$model))  
  (1-exp((rr$dev-rr$null)/n))/(1-exp(-rr$null/n))  
}  
  
require(MASS)  
table(bacteria$y,bacteria$week)  
table(bacteria$y,bacteria$trt)  
res <- glm(y ~ week+trt, family=binomial, data=bacteria)  
NagelkerkeR2(res)  
summary(res)  
exp(coef(res))  
exp(confint(res))
```

この課題で立てるモデルは、菌の検出の有無を従属変数、週数と処置を独立変数とするロジスティック回帰である。週数は数値のまま入れ、影響を調整すべき共変量として扱えばよい。処置は3水準あるが、プラセボをリファレンスにしたときにコンプライアンスの悪い服薬群とコンプライアンスの良い服薬群で<sup>\*1</sup>、それぞれどれくらい菌の検出が減るか（オッズ比がいくつになるか）を明らかにすることが目的である。ロジスティック回帰に先立ち、菌の検出の有無と週数、処置をそれぞれクロス集計した結果は次の通りである。

```
y / week  
  0  2  4  6 11  
n  5  4 11 11 12  
y 45 40 31 29 32  
y / trt  
  placebo drug drug+  
n      12  18  13  
y      84  44  49
```

0週、2週では大半の子供に菌が検出されていたが4週以降になると検出されない子供が増えてくることがわかる。また、プラセボ群では大半の子供に菌が検出されるが、投薬群では菌が検出されない子供の割合が増えてきているようにみえる。

ロジスティック回帰分析の結果を次の表にまとめる。

表. アボリジニの中耳炎乳児へのアモキシリン投与が *H. influenzae* の有無に与える影響のロジスティック回帰分析\*

独立変数	オッズ比	95%信頼区間		p 値
		下限	上限	
プラセボ	1	—	—	—
アモキシリン投与, コンプライアンス低	0.331	0.140	0.752	0.009
アモキシリン投与, コンプライアンス高	0.521	0.215	1.253	0.144

Nagelkerke の  $R^2$ : 0.095, AIC: 211.8,  $D_{null}$ : 217.4 (自由度 219),  $D$ : 203.8 (自由度 216)

\* 投与週数の効果を共変量として調整した (偏回帰係数  $-0.116$ ,  $p = 0.009$ )

アモキシリン投与かつコンプライアンスが低かった群では菌検出のオッズ比が有意に1より小さく、約1/3に減少していた。コンプライアンスが高かった群でも菌検出のオッズ比は1より小さいが有意水準5%で有意ではなかった。コンプライアンスが低い群の方が抗生物質の効きが良いと考えると不思議な結果である。

他の情報がないのでスペキュレーションしかできないが、コンプライアンスが低い群の方が効くのではなく、良く効いたために少ししか飲まなくても症状が軽快し、以後飲まなくなった、つまり効いた人の方がコンプライアンスが悪くなった可能性があるかもしれない。

\*1 本当は、コンプライアンスの良い服薬群とコンプライアンスの悪い服薬群などというものができてしまった時点で、RCTとしては問題がある。