一般化線形モデル（概論）

1 概要

一般化線形モデル（Generalized Linear Model; GLM）は線形モデルの拡張として提案されたものである。線形回帰モデル，対数線形モデル，ロジスティック回帰モデルなどを統一的に表現できるものである。一般化線形モデルは次の 3 つの成分から規定されるモデルである。

1. **ランダム成分（random conponent）**
   説明変数が与えられたときの目的変数は，（正準形をもつ）指数分布族に従うと仮定される。指数分布族には，正規分布，二項分布，ポアソン分布などがある。
2. **系統的成分（systematic component）**
   説明変数は線形的にモデルに関与するとして，線形予測子を説明変数の線形結合として次のように定義する。
   $$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}$$
3. **連結関数（link function）**
   線形予測子は目的変数の平均 $\mu_i = E[Y_i]$ の関数
   $$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}$$
   と仮定される。線形予測子と平均の関係を規定する関数 $g(\cdot)$ は，連結関数とよばれ，通常滑らかで単調性をもつと仮定する。

2 指数分布族

単一のパラメータ θ をもつ確率分布に従う一つの確率変数 Y について考える。Y の確率（密度）関数が次式で表すことができるとき，Y は指数分布族に属するという。

$$f(y;\theta) = s(y)t(\theta)\exp(a(y)b(\theta)) \tag{1}$$

ただし，a, b, s, t は既知の関数とする。y と θ の間には，対称的な関係があることに注意する。このことは，(1) 式を次式のように表現するとより明らかになる。

$$f(y;\theta) = \exp[a(y)b(\theta) + c(\theta) + d(y)] \tag{2}$$

ただし，$s(y) = \exp(d(y))$, $t(\theta) = \exp(c(\theta))$ である。

$a(y) = y$ のとき，その分布は**正準形**（canonical form）であり，$b(\theta)$ は分布の**自然母数**（natural parameter）と呼ばれる。もし，関心のある母数 θ 以外に他の母数があるとき，関数 a, b, c, d を構成する**局外母数**（nuisance parameter）とされ，既知として扱われる。

代表的な確率分布は，指数分布族に属している。例えば，正規分布，ポアソン分布，二項分布は，すべて正準形として記述できる。実際に，このことを確認してみる。

2.1 正規分布

正規分布 $N(\mu, \sigma^2)$ の確率密度関数は，次のように表される。

$$\begin{aligned} f(y;\mu,\sigma^2) &= \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \\ &= \exp\left[-\frac{y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right] \\ &= \exp\left[\frac{1}{\sigma^2}\left(-\frac{y^2}{2} + y\mu\right) - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right] \end{aligned}$$

まず，μ を関心のある母数，σ^2 を局外母数とした場合を考える。ここで，

$$a(y) = y, \quad b(\mu) = \frac{\mu}{\sigma^2}, \quad c(\mu) = -\frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2), \quad d(y) = -\frac{y^2}{2\sigma^2}$$

または

$$a(y) = y, \quad b(\mu) = \frac{\mu}{\sigma^2}, \quad c(\mu) = -\frac{\mu^2}{2\sigma^2}, \quad d(y) = -\frac{y^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)$$

とすれば，(2) 式を満たすことがわかる。したがって，μ を関心のある母数，σ^2 を局外母数とした正規分布は，指数分布族に属している。さらに，$a(y) = y$ であるので，正準形である。

次に，σ^2 を関心のある母数，μ を局外母数とした場合を考える。ここで，

$$a(y) = -\frac{y^2}{2} + y\mu, \quad b(\sigma^2) = \frac{1}{\sigma^2}, \quad c(\sigma^2) = -\frac{\mu^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2), \quad d(y) = 0$$

とすれば，(2) 式を満たすことがわかる。したがって，σ^2 を関心のある母数，μ を局外母数とした正規分布は，指数分布族に属している。さらに，$a(y) \neq y$ であるので，正準形ではない。

2.2 ポアソン分布

ポアソン分布は，計数データのモデルとしてよく用いられる。典型的には，非常に短い時間の間隔，あるいは空間の間隔において事象が起こる確率が非常に小さく，それらの事象が独立に起こるという場合に，一定の時間や空間において発生する事象数の確率モデルのような場合である。例えば，ある交差点で 1 日当たりに事故が起こる回数，あるコンビニに 1 日当たりに来る客数，ある地域で 1 日当たりの新型コロナウイルス感染者数，などがある。

ポアソン分布 $Po(\mu)$ の確率関数は，次のように表される。

$$\begin{aligned} f(y;\mu) &= \frac{\mu^y}{y!}\exp(-\mu) \\ &= \exp\left[\log\left(\frac{\mu^y}{y!}\right) - \mu\right] \\ &= \exp(y\log\mu - \log y! - \mu) \end{aligned}$$

ここで，

$$a(y) = y, \quad b(\mu) = \log\mu, \quad c(\mu) = -\mu, \quad d(y) = -\log y!$$

とすれば，(2) 式を満たすことがわかる。したがって，ポアソン分布は，指数分布族に属している。さらに，$a(y) = y$ であるので，正準形である。

2.3 二項分布

二項分布は，二値の出力をもつ観測値の系列に対して，第一選択として考えられるモデルである。例えば，ある治療を受けた患者の中で有効であった患者数（一人ずつについて起こり得る事象は，有効か無効かのいずれかとする）。

π を関心のある母数，n は既知と仮定し，二項分布 $Bi(n, \pi)$ の確率関数は，次のように表される。

$$\begin{aligned} f(y;\pi) &= \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \exp\left[\log\binom{n}{y} + y\log\pi + (n-y)\log(1-\pi)\right] \\ &= \exp\left[\log\binom{n}{y} + y\log\left(\frac{\pi}{1-\pi}\right) + n\log(1-\pi)\right] \end{aligned}$$

ここで，

$$a(y) = y, \quad b(\pi) = \log\left(\frac{\pi}{1-\pi}\right), \quad c(\pi) = n\log(1-\pi), \quad d(y) = \log\binom{n}{y}$$

とすれば，(2) 式を満たすことがわかる。したがって，二項分布は，指数分布族に属している。さらに，$a(y) = y$ であるので，正準形である。

3 指数分布族の性質

$a(Y)$ の期待値と分散を求めることを考える。確率変数 Y が連続型か離散型かにより導出方法がことなるため，それぞれの場合について述べる。

3.1 連続型の確率変数

確率変数 Y が連続型の確率変数であれば，積分と微分の順序を入れ替えることができるという条件のもとで，任意の確率密度関数について成り立つ次式の結果を用いる。

$$\int f(y;\theta)dy = 1 \tag{3}$$

ただし，積分の範囲は，y の取りうる範囲の全体とする。

θ に関して (3) 式の両辺を微分すると，次式を得る。

$$\frac{\partial}{\partial\theta}\int f(y;\theta)dy = \frac{\partial}{\partial\theta}1 = 0 \tag{4}$$

積分と微分の順序を入れ替えると，(4) 式は次のように表すことができる。

$$\int \frac{\partial}{\partial\theta} f(y;\theta)dy = 0 \tag{5}$$

同様に，(3) 式を θ に関して二回微分し，積分と微分の順序を入れ替えると次式を得る。

$$\frac{\partial^2}{\partial\theta^2}\int f(y;\theta)dy = \int \frac{\partial^2}{\partial\theta^2} f(y;\theta)dy = 0 \tag{6}$$

これらの結果を指数分布族に属する分布に適用する。(2) 式より，指数分布族に属する分布の確率密度関数は

$$f(y;\theta) = \exp[a(y)b(\theta) + c(\theta) + d(y)]$$

であることから，

$$\frac{\partial}{\partial\theta} f(y;\theta) = [a(y)b'(\theta) + c'(\theta)]f(y;\theta) \quad \left(\because b'(\theta) = \frac{\partial b(\theta)}{\partial\theta}, \quad c'(\theta) = \frac{\partial c(\theta)}{\partial\theta}\right)$$

となる。(5) 式より，次式を得る。

$$\begin{aligned}
&\int [a(y)b'(\theta) + c'(\theta)]f(y;\theta)dy = 0 \\
&\Leftrightarrow b'(\theta)\int a(y)f(y;\theta)dy + c'(\theta)\int f(y;\theta)dy = 0 \\
&\Leftrightarrow b'(\theta)E[a(Y)] + c'(\theta) = 0 \quad \left(\because E[a(Y)] = \int a(y)f(y;\theta)dy, \quad \int f(y;\theta)dy = 1\right) \\
&\Leftrightarrow E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}
\end{aligned}$$

同様に，$V[a(Y)]$ を求めることができる。

$$\frac{\partial^2}{\partial\theta^2} f(y;\theta) = [a(y)b''(\theta) + c''(\theta)]f(y;\theta) + [a(y)b'(\theta) + c'(\theta)]^2 f(y;\theta) \tag{7}$$
$$\left(\because b''(\theta) = \frac{\partial^2 b(\theta)}{\partial\theta^2}, \quad c''(\theta) = \frac{\partial^2 c(\theta)}{\partial\theta^2}\right)$$

(7) 式の右辺第二項は，次式となる。

$$\begin{aligned}
[a(y)b'(\theta) + c'(\theta)]^2 f(y;\theta) &= [(a(y)b'(\theta))^2 + 2a(y)b'(\theta)c'(\theta) + (c'(\theta))^2]f(y;\theta) \\
&= [(a(y)b'(\theta))^2 + 2a(y)b'(\theta)(-b'(\theta)E[a(Y)]) + (-b'(\theta)E[a(Y)])^2]f(y;\theta) \\
&\qquad (\because c'(\theta) = -b'(\theta)E[a(Y)]) \\
&= (b'(\theta))^2[a(y) - E[a(Y)]]^2 f(y;\theta)
\end{aligned}$$

したがって，(7) 式は次式のように表せる。

$$\begin{aligned}
\int \frac{\partial^2}{\partial\theta^2} f(y;\theta)dy &= \int [a(y)b''(\theta) + c''(\theta)]f(y;\theta)dy + \int (b'(\theta))^2[a(y) - E[a(Y)]]^2 f(y;\theta)dy \\
&= b''(\theta)E[a(Y)] + c''(\theta) + (b'(\theta))^2 V[a(Y)] \\
&\qquad \left(\because V[a(Y)] = \int [a(y) - E[a(Y)]]^2 f(y;\theta)dy\right) \\
&= 0
\end{aligned}$$

この式を整理すると

$$\begin{aligned}
V[a(Y)] &= -\frac{b''(\theta)E[a(Y)] + c''(\theta)}{(b'(\theta))^2} \\
&= -\frac{b''(\theta)\left(-\dfrac{c'(\theta)}{b'(\theta)}\right) + c''(\theta)}{(b'(\theta))^2} \quad \left(\because E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}\right) \\
&= \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3}
\end{aligned}$$

以上の結果を整理すると，指数分布族における分布の期待値や分散は，次式で表されることがわかる。

$$E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}, \quad V[a(Y)] = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3}$$

3.2 離散型の確率変数

確率変数 Y が離散型の確率変数であれば，任意の確率関数について成り立つ次式の結果を用いる。

$$\sum f(y;\theta) = 1 \tag{8}$$

ただし，和の範囲は，y の取りうる値のすべてとする。θ に関して (8) 式の両辺を微分すると，次式を得る。

$$\frac{\partial}{\partial\theta}\sum f(y;\theta) = \frac{\partial}{\partial\theta}1 = 0 \Leftrightarrow \sum \frac{\partial}{\partial\theta} f(y;\theta) = 0 \tag{9}$$

同様に，(8) 式を θ に関して二回微分すると，次式を得る。

$$\frac{\partial^2}{\partial\theta^2}\sum f(y;\theta) = \sum \frac{\partial^2}{\partial\theta^2} f(y;\theta) = 0 \tag{10}$$

これらの結果を指数分布族に属する分布に適用する。(2) 式より，指数分布族に属する分布の確率密度関数は

$$f(y;\theta) = \exp[a(y)b(\theta) + c(\theta) + d(y)]$$

であることから，

$$\frac{\partial}{\partial\theta} f(y;\theta) = [a(y)b'(\theta) + c'(\theta)]f(y;\theta) \quad \left(\because b'(\theta) = \frac{\partial b(\theta)}{\partial\theta}, \quad c'(\theta) = \frac{\partial c(\theta)}{\partial\theta}\right)$$

となる。(9) 式より，次式を得る。

$$\begin{aligned}
&\sum [a(y)b'(\theta) + c'(\theta)]f(y;\theta) = 0 \\
&\Leftrightarrow b'(\theta)\sum a(y)f(y;\theta) + c'(\theta)\sum f(y;\theta) = 0 \\
&\Leftrightarrow b'(\theta)E[a(Y)] + c'(\theta) = 0 \quad \left(\because E[a(Y)] = \sum a(y)f(y;\theta), \quad \sum f(y;\theta) = 1\right) \\
&\Leftrightarrow E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}
\end{aligned}$$

同様に，$V[a(Y)]$ を求めることができる。

$$\frac{\partial^2}{\partial\theta^2} f(y;\theta) = [a(y)b''(\theta) + c''(\theta)]f(y;\theta) + [a(y)b'(\theta) + c'(\theta)]^2 f(y;\theta) \tag{11}$$
$$\left(\because b''(\theta) = \frac{\partial^2 b(\theta)}{\partial\theta^2}, \quad c''(\theta) = \frac{\partial^2 c(\theta)}{\partial\theta^2}\right)$$

(11) 式の右辺第二項は，次式となる。

$$\begin{aligned}
[a(y)b'(\theta) + c'(\theta)]^2 f(y;\theta) &= [(a(y)b'(\theta))^2 + 2a(y)b'(\theta)c'(\theta) + (c'(\theta))^2]f(y;\theta) \\
&= [(a(y)b'(\theta))^2 + 2a(y)b'(\theta)(-b'(\theta)E[a(Y)]) + (-b'(\theta)E[a(Y)])^2]f(y;\theta) \\
&\qquad (\because c'(\theta) = -b'(\theta)E[a(Y)]) \\
&= (b'(\theta))^2[a(y) - E[a(Y)]]^2 f(y;\theta)
\end{aligned}$$

したがって，(10) 式は次式のように表せる。

$$\begin{aligned}\sum \frac{\partial^2}{\partial\theta^2} f(y;\theta)dy &= \sum [a(y)b''(\theta) + c''(\theta)]f(y;\theta) + \sum (b'(\theta))^2[a(y) - E[a(Y)]]^2 f(y;\theta) \\ &= b''(\theta)E[a(Y)] + c''(\theta) + (b'(\theta))^2 V[a(Y)] \\ &\qquad \left(\because V[a(Y)] = \sum [a(y) - E[a(Y)]]^2 f(y;\theta)\right) \\ &= 0\end{aligned}$$

この式を整理すると

$$\begin{aligned}V[a(Y)] &= -\frac{b''(\theta)E[a(Y)] + c''(\theta)}{(b'(\theta))^2} \\ &= -\frac{b''(\theta)\left(-\dfrac{c'(\theta)}{b'(\theta)}\right) + c''(\theta)}{(b'(\theta))^2} \qquad \left(\because E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}\right) \\ &= \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3}\end{aligned}$$

以上の結果を整理すると，指数分布族における分布の期待値や分散は，次式で表されることがわかる。

$$E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}, \quad V[a(Y)] = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3}$$

問 1

指数分布族の性質を用いて，正規分布，ポアソン分布，二項分布の期待と分散を求めよ。ただし，正規分布は，μ を関心のある母数，σ^2 を局外母数とした場合，二項分布は π を関心のある母数，n は既知とした場合とする。

3.3 スコア統計量

対数尤度関数の導関数の期待値と分散を求めることを考える。(2) 式から，指数分布族に属する分布の対数尤度関数は，次式のようになる。

$$l(\theta; y) = a(y)b(\theta) + c(\theta) + d(y)$$

対数尤度関数の θ に関する導関数は，次式のようになる。

$$U(\theta; y) = \frac{\partial l(\theta; y)}{\partial\theta} = a(y)b'(\theta) + c'(\theta)$$

関数 U は**スコア統計量**（score statistic）とよばれる。U は y に依存しているので確率変数である。したがって，次式のように表すことができる。

$$U = a(Y)b'(\theta) + c'(\theta)$$

確率変数 U の期待値は，次式のようになる。

$$\begin{aligned} E[U] &= E[a(Y)b'(\theta) + c'(\theta)] \\ &= b'(\theta)E[a(Y)] + c'(\theta) \\ &= b'(\theta)\left(-\frac{c'(\theta)}{b'(\theta)}\right) + c'(\theta) \quad \left(\because E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)}\right) \\ &= 0 \end{aligned}$$

確率変数 U の分散は，**情報量**（information）とよばれ，$\mathfrak{I}$ と表す。確率変数の線形変換の分散を計算する公式を用いると，次式を得る。

$$\begin{aligned} \mathfrak{I} = V[U] &= V[a(Y)b'(\theta) + c'(\theta)] \\ &= (b'(\theta))^2 V[a(Y)] \\ &= (b'(\theta))^2 \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3} \quad \left(\because V[a(Y)] = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{(b'(\theta))^3}\right) \\ &= \frac{b''(\theta)c'(\theta)}{b'(\theta)} - c''(\theta) \end{aligned}$$

スコア統計量 U は，一般化線形モデルにおけるパラメータ推測に利用される。

問 2

次式が成り立つことを示せ。ただし，$U' = \dfrac{\partial U}{\partial \theta}$ とする。

$$V[U] = E[U^2] = -E[U']$$

参考文献

[1] 江金芳．(2016)．一般化線形モデル．朝倉書店.

[2] Dobson, A. J and Barnett, A. G. (2018). *An introduction to generalized linear models, 4th Edition.* Chapman and Hall/CRC.