

# 最尤法について

????,050610, 1111,18,121003

## 1 最尤解

$Y_i, i = 1, 2, \dots, n$  を独立な観測値とする。 $f(y_i|\theta)$  をデータの確率関数（密度関数）とする。データの同時分布の確率関数は

$$f(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^n f(y_i|\theta) \quad (1)$$

である。これは、母数の値が  $\theta$  の時に、 $y_1, y_2, \dots, y_n$  というデータを観測する確率（密度）であり、その値が高いほど、それらのデータ系列（パターン）が出現する確率が高い。

この同時確率関数を母数  $\theta$  の関数と見たものが尤度関数である。

$$L(\theta) = f(y_1, y_2, \dots, y_n|\theta) \quad (2)$$

これは、データ系列を与えられた時に、母数の値がどの程度尤もらしいかを表す値と解釈することが出来る。ただし、 $\theta$  の確率を与えるものではない。（ $\theta$  は確率変数ではないので確率は付与できない。）

なお、パラメタを含まない部分は、尤度関数の定義から落としても良い。その意味で、

$$L(\theta) \propto f(y_1, y_2, \dots, y_n|\theta) \quad (3)$$

と書く場合がある。

尤度関数の対数を対数尤度関数という。

$$\ln L(\theta) = \log L(\theta) = \sum_{i=1}^n \log f(y_i|\theta) \quad (4)$$

尤度関数もしくは対数尤度関数の最大値を与える母数  $\theta$  の値、 $\hat{\theta}$ 、をその母数の推定値とする方法を、最尤解と呼ぶ。最尤解を求めるためには対数尤度関数の微分を 0 とおいた方程式

$$d \ln L(\theta) = \frac{d \ln L(\theta)}{d \theta} = 0 \quad (5)$$

を解けばよい。なお、得られた解で、対数尤度の 2 次微分が負になっていることが必要である。

**正規分布の場合（ $\sigma = 1$  を仮定）**

$$f(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - \theta)^2\right) \quad (6)$$

$$L(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - \theta)^2\right) \quad (7)$$

$$\propto \prod_{i=1}^n \exp\left(-\frac{1}{2}(y_i - \theta)^2\right) \quad (8)$$

$$\ln L(\theta) = \sum_{i=1}^n -\frac{1}{2}(y_i - \theta)^2 + \text{const} \quad (9)$$

$$d \ln L(\theta) = \sum_{i=1}^n (y_i - \theta) = 0 \quad (10)$$

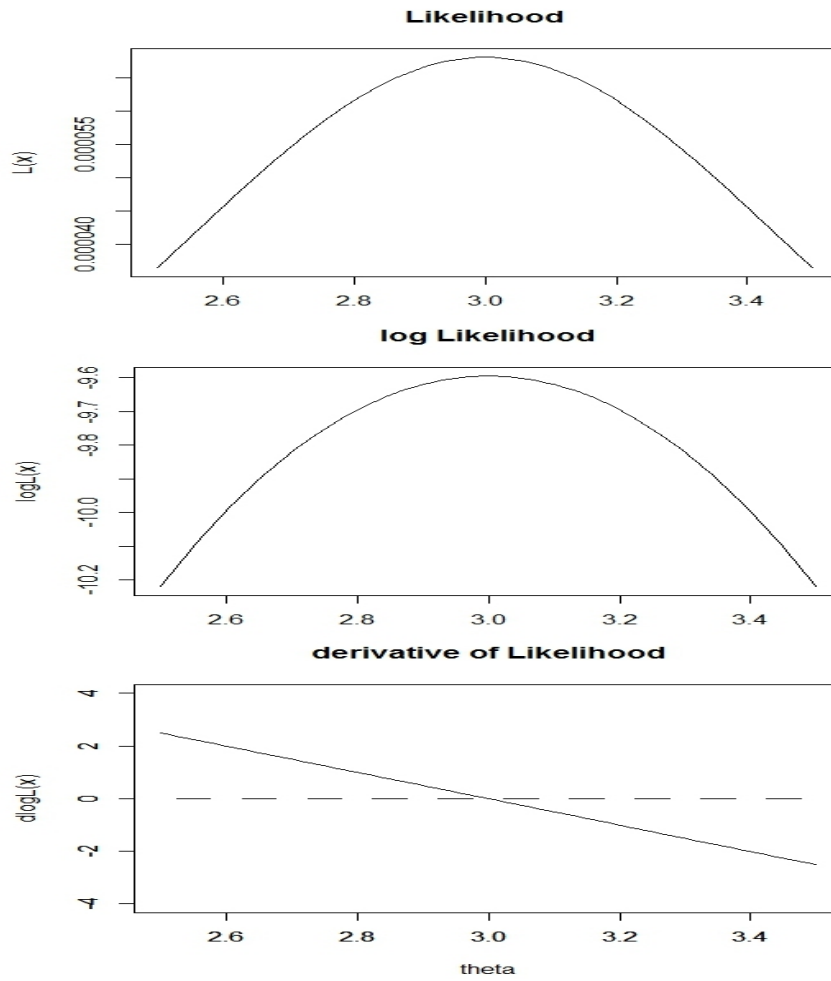
したがって、

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i \quad (11)$$

例（正規分布）

$y = \{1, 2, 3, 4, 5\}$  の場合

THETA	L	LL	DLL
2.5	0.000019	-10.8236	2.5000
2.6	0.000025	-10.5986	2.0000
2.7	0.000029	-10.4236	1.5000
2.8	0.000033	-10.2986	1.0000
2.9	0.000036	-10.2236	0.5000
3	0.000037	-10.1986	0.0000 *
3.1	0.000036	-10.2236	0.5000
3.2	0.000033	-10.2986	1.0000
3.3	0.000029	-10.4236	1.5000
3.4	0.000025	-10.5986	2.0000
3.5	0.000019	-10.8236	2.5000



ベルヌーイ分布の場合

$$f(y_1, y_2, \dots, y_n | \theta) = \prod_{i=1}^n \theta^{y_i} (1 - \theta)^{1-y_i} \quad (12)$$

$$L(\theta) = \prod_{i=1}^n \theta^{y_i} (1 - \theta)^{1-y_i} \quad (13)$$

$$\ln L(\theta) = \sum_{i=1}^n (y_i \log \theta + (1 - y_i) \log(1 - \theta)) \quad (14)$$

$$d \ln L(\theta) = \sum_{i=1}^n \left( \frac{y_i}{\theta} - \frac{(1 - y_i)}{(1 - \theta)} \right) = 0 \quad (15)$$

したがって、

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i \quad (16)$$

例 (2項分布)

$\mathbf{y} = \{0, 1, 1, 0, 1\}$  の場合

THETA	L	LL	DLL
0.1	0.0008	-7.1185	27.7778
0.2	0.0051	-5.2746	12.5000
0.3	0.0132	-4.3253	7.1429
0.4	0.0230	-3.7705	4.1667
0.5	0.0313	-3.4657	2.0000
0.6	0.0346	-3.3651	-0.0000 *
0.7	0.0309	-3.4780	-2.3810
0.8	0.0205	-3.8883	-6.2500
0.9	0.0073	-4.9213	-16.6667

正規分布の場合

$$f(y_1, y_2, \dots, y_n | \mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}(y_i - \mu)^2/\sigma^2\right) \quad (17)$$

$$L(\mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}(y_i - \mu)^2/\sigma^2\right) \quad (18)$$

$$\propto \left(\frac{1}{\sigma}\right)^n \prod_{i=1}^n \exp\left(-\frac{1}{2}(y_i - \mu)^2/\sigma^2\right) \quad (19)$$

$$\ln L(\mu, \sigma) = -n \ln \sigma + \sum_{i=1}^n -\frac{1}{2}(y_i - \mu)^2/\sigma^2 + \text{const} \quad (20)$$

$$= -\frac{1}{2} \left( n \ln \sigma^2 + \sum_{i=1}^n (y_i - \mu)^2/\sigma^2 \right) + \text{const} \quad (21)$$

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n (y_i - \mu)/\sigma^2 \quad (22)$$

$$\frac{\partial \ln L}{\partial \sigma} = \quad (23)$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \quad (24)$$

$$\frac{\partial \ln L}{\partial \tau} = \quad \tau = \frac{1}{\sigma^2} \quad (25)$$

$$(26)$$

したがって、

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n y_i \quad (27)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\mu})^2 \quad (28)$$

## 2 回帰分析

### 2.1 正規分布の場合：回帰分析

データ

$$(x_i, y_i), i = 1, 2, \dots, n \quad \text{ただし } y_i \text{ は連続量} \quad (29)$$

モデル

$$y_i = \alpha x_i + \beta + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (30)$$

または

$$\mathcal{E}(y_i | x_i) = \hat{y}_i = \alpha x_i + \beta \quad (31)$$

モデルのパラメタは  $\xi = \{\alpha, \beta, (\sigma^2)\}$

尤度関数

$$L(\xi) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}(y_i - (\alpha x_i + \beta))^2 / \sigma^2\right) \quad (32)$$

対数尤度関数

$$\ln L(\xi) = -\frac{1}{2} \left( n \ln(\sigma^2) + \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - (\alpha x_i + \beta))^2 \right) + \text{const} \quad (33)$$

これを最大にするように  $\xi$  を求める。すなわち、連立方程式

$$\frac{\partial \ln L}{\partial \alpha} = 0, \quad \frac{\partial \ln L}{\partial \beta} = 0, \quad \frac{\partial \ln L}{\partial \sigma} = 0 \quad (34)$$

を解く。答えは

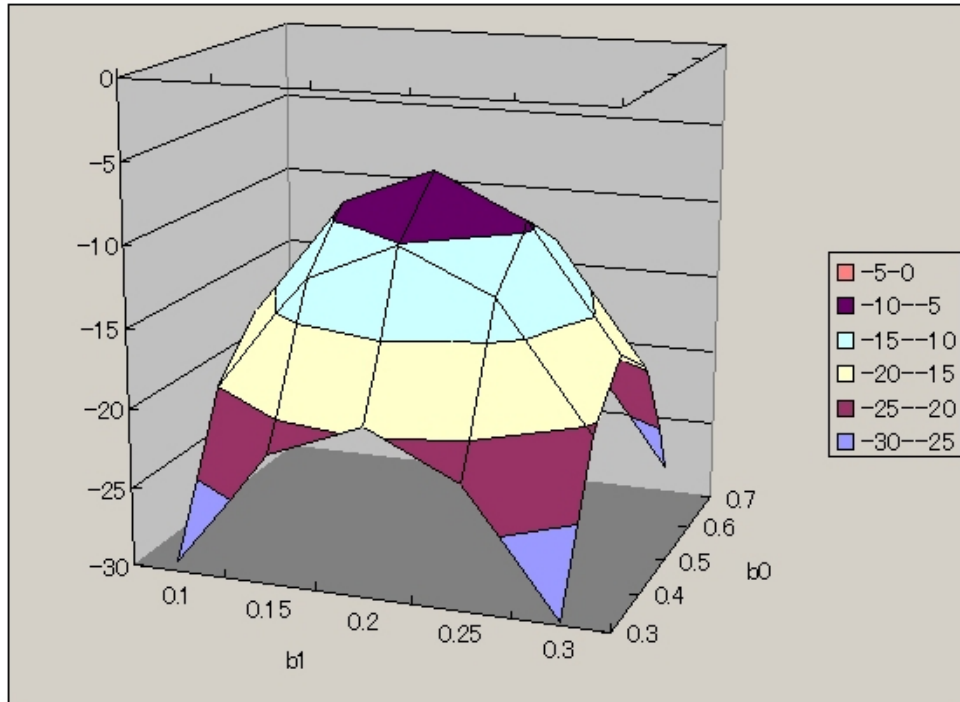
$$\hat{\alpha} = \frac{\text{cov}(x, y)}{\text{var}(x)}, \quad \hat{\beta} = \bar{y} - \hat{\alpha} \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - (\hat{\alpha} x_i + \hat{\beta}))^2 \quad (35)$$

Data1				最尤解の探索		
Obs	X	y	yhat	alpha	beta	LL
				0.3	0.1	-29.316
1	-3.0	0.000	-0.200	0.3	0.15	-21.841
2	-2.6	0.000	-0.080	0.3	0.2	-19.438
3	-2.3	0.000	0.010	0.3	0.25	-22.105
4	-1.9	0.000	0.130	0.3	0.3	-29.843
5	-1.5	0.200	0.250	0.4	0.1	-19.984
6	-1.1	0.166	0.370	0.4	0.15	-12.509
7	-0.8	0.285	0.460	0.4	0.2	-10.106
8	-0.4	0.125	0.580	0.4	0.25	-12.773
9	0.0	0.777	0.700	0.4	0.3	-20.511
10	0.4	0.875	0.820	0.5	0.1	-16.614
11	0.8	0.714	0.940	0.5	0.15	-9.139
12	1.1	1.000	1.030	0.5	0.2	-6.735 *
13	1.5	0.800	1.150	0.5	0.25	-9.402
14	1.9	1.000	1.270	0.5	0.3	-17.141
15	2.3	0.666	1.390	0.6	0.1	-19.205
16	2.6	1.000	1.480	0.6	0.15	-11.730
17	3.0	1.000	1.600	0.6	0.2	-9.326
				0.6	0.25	-11.994
				0.6	0.3	-19.732
				0.7	0.1	-27.757
				0.7	0.15	-20.283
				0.7	0.2	-17.879
				0.7	0.25	-20.546
				0.7	0.3	-28.284

**最尤解**

$\beta$  0.5065  
 $\alpha$  0.1987  
 $\sigma$  0.1689

ただし、表中 yhat は上記の最尤解で計算したもの。



## 2.2 2項分布の場合：ロジスティック回帰分析

データ

$$(x_i, y_i), i = 1, 2, \dots, n \quad \text{ただし } y_i \text{ は } 0 \text{ か } 1 \quad (36)$$

これを  $x_i$  の異なる値ごとに度数分布の形にしたもの

$$(x_i, f_i, r_i), i = 1, 2, \dots, N \quad \text{ただし } f_i \text{ は試行数 } r_i \text{ は成功数} \quad (37)$$

モデル

確率は  $(0, 1)$  の区間に入らなければならないので、

$$\Pr(Y_i = 1|x_i) = \frac{1}{1 + \exp(-(\alpha x_i + \beta))} = P(x_i|\alpha, \beta) = \pi_i \quad (38)$$

とする。これは

$$\log \frac{\Pr(Y_i = 1|x_i)}{1 - \Pr(Y_i = 1|x_i)} = \alpha x_i + \beta \quad (39)$$

と同値である。なお、モデルのパラメタは  $\xi = \{\alpha, \beta\}$  であり、

$$\mathcal{E}(Y_i|x_i) = P(x_i|\alpha, \beta) = \pi_i \quad (40)$$

$$\text{var}(Y_i|x_i) = \pi_i(1 - \pi_i) \quad (41)$$

となる。また、 $(f_i, r_i)$  を試行数と度数として、

$$\mathcal{E}(R_i|x_i) = f_i \times P(x_i|\alpha, \beta) = f_i \pi_i \quad (42)$$

$$\text{var}(R_i|x_i) = f_i \pi_i(1 - \pi_i) \quad (43)$$



である。

尤度関数

$$L(\boldsymbol{\xi}) = \prod_{i=1}^n P(x_i|\alpha, \beta)^{y_i} (1 - P(x_i|\alpha, \beta))^{1-y_i} \quad (44)$$

$$L(\boldsymbol{\xi}) = \prod_{i=1}^n P(x_i|\alpha, \beta)^{r_i} (1 - P(x_i|\alpha, \beta))^{f_i-r_i} \quad (45)$$

対数尤度関数

$$\ln L(\boldsymbol{\xi}) = \sum_{i=1}^n y_i \log P(x_i|\alpha, \beta) + (1 - y_i) \log(1 - P(x_i|\alpha, \beta)) \quad (46)$$

$$\ln L(\boldsymbol{\xi}) = \sum_{i=1}^n r_i \log P(x_i|\alpha, \beta) + (f_i - r_i) \log(1 - P(x_i|\alpha, \beta)) \quad (47)$$

これを最大にするように  $x_i$  を求める。すなわち、連立方程式

$$\frac{\partial \ln L}{\partial \alpha} = 0, \quad \frac{\partial \ln L}{\partial \beta} = 0, \quad (48)$$

を解く。答えは解析的には解けない。

IRT のパラメタとの関係

指数関数の中が、IRT の場合は  $-1.7a(x - b)$ , LR:  $-(\alpha x + \beta)$  であるから、 $a = \alpha/1.7$ ,  $b = -\beta/\alpha$  となる。

例

Data1						最尤解の探索		
Obs	X	F	R	P	phat	ALPHA	BETA	LL
						1.0	-0.5	-40.6408
1	-3.0	1	0	0.000	0.015	1.0	-0.3	-38.9813
2	-2.6	2	0	0.000	0.026	1.0	-0.1	-37.8991
3	-2.3	3	0	0.000	0.040	1.0	0.1	-37.3991
4	-1.9	4	0	0.000	0.071	1.0	0.3	-37.4813
5	-1.5	5	1	0.200	0.120	1.0	0.5	-38.1408
6	-1.1	6	1	0.166	0.198	1.0	0.7	-39.3681
7	-0.8	7	2	0.285	0.277	1.0	0.9	-41.1493
8	-0.4	8	1	0.125	0.407	1.2	-0.5	-39.6113
9	0.0	9	7	0.777	0.552	1.2	-0.3	-38.0549
10	0.4	8	7	0.875	0.689	1.2	-0.1	-37.0249
11	0.8	7	5	0.714	0.799	1.2	0.1	-36.5249
12	1.1	6	6	1.000	0.860	1.2	0.3	-36.5549
13	1.5	5	4	0.800	0.917	1.2	0.5	-37.1113
14	1.9	4	4	1.000	0.952	1.2	0.7	-38.1873
15	2.3	3	2	0.666	0.972	1.2	0.9	-39.7725
16	2.6	2	2	1.000	0.982	1.4	-0.5	-39.1354
17	3.0	1	1	1.000	0.990	1.4	-0.3	-37.6707
						1.4	-0.1	-36.6870
						1.4	0.1	-36.1870
						1.4	0.3	-36.1707 *
						1.4	0.5	-36.6354
						1.4	0.7	-37.5759
						1.4	0.9	-38.9847
						1.6	-0.5	-39.0962
						1.6	-0.3	-37.7122
						1.6	-0.1	-36.7692
						1.6	0.1	-36.2692
						1.6	0.3	-36.2122
						1.6	0.5	-36.5962
						1.6	0.7	-37.4173
						1.6	0.9	-38.6698
						1.8	-0.5	-39.4032
						1.8	-0.3	-38.0898
						1.8	-0.1	-37.1823
						1.8	0.1	-36.6823
						1.8	0.3	-36.5898
						1.8	0.5	-36.9032
						1.8	0.7	-37.6195
						1.8	0.9	-38.7345
						2.0	-0.5	-39.9865
						2.0	-0.3	-38.7348
						2.0	-0.1	-37.8585
						2.0	0.1	-37.3585

ただし、 phat は最尤推定値で計算したもの。

2.0	0.3	-37.2348
2.0	0.5	-37.4865
2.0	0.7	-38.1110
2.0	0.9	-39.1051

最优解

$\alpha$  1.4642 (0.964)

$\beta$

