

# 第5回 母数の推定 最尤解

## 最尤法

$Y_i, i = 1, 2, \dots, n$  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。

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

多くの場合、対数尤度関数の方が尤度関数よりも簡単な形をしている。

## 離散の場合の例

### ■ 5点満点のテストの真値の推定

- ◆  $E \sim N(0,1)$
- ◆  $T = 2.0, 2.5, 3.0, 3.5, 4.0$
- ◆  $X = T + E,$

★  $X | T \sim N(T, 1) \quad f(X = x | T = t) \propto \exp(-0.5(x - t)^2)$

### ■ T を与えたときの X の確率分布（縦に足して 1）

n=1 回の観測値から

	t=1.0	t=1.2	t=1.5	t=1.8	t=2.0	t=2.2	t=2.5	t=2.8	t=3.0	t=3.2	t=3.5	t=3.8	t=4.0
x=0	0.26	0.19	0.13	0.09	0.05	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00
x=1	0.42	0.40	0.36	0.30	0.24	0.18	0.13	0.09	0.05	0.03	0.02	0.01	0.00
x=2	0.26	0.31	0.36	0.39	0.40	0.39	0.35	0.30	0.24	0.18	0.13	0.09	0.06
x=3	0.06	0.09	0.13	0.18	0.24	0.30	0.35	0.39	0.40	0.39	0.36	0.31	0.26
x=4	0.00	0.01	0.02	0.03	0.05	0.09	0.13	0.18	0.24	0.30	0.36	0.40	0.42
x=5	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.05	0.09	0.13	0.19	0.26

$$f(X_1 = x_1 \& X_2 = x_2 | T = t) \propto$$

- 独立に2回受けた場合

$$\exp(-0.5(x_1 - t)^2) \times \exp(-0.5(x_2 - t)^2)$$

	2.0	2.5	3.0	3.5	4.0
00	0.0029	0.0003	0	0	0
01	0.0132	0.0023	0.0002	0	0
02	0.0217	0.0062	0.0011	0.0001	0
03	0.0132	0.0062	0.0018	0.0003	0
04	0.0029	0.0023	0.0011	0.0003	0.0001
05	0.0002	0.0003	0.0002	0.0001	0
10	0.0132	0.0023	0.0002	0	0
11	0.0591	0.0168	0.0029	0.0003	0
12	0.0974	0.0458	0.0132	0.0024	0.0003
13	0.0591	0.0458	0.0217	0.0064	0.0012
14	0.0132	0.0168	0.0132	0.0064	0.002
15	0.0011	0.0023	0.0029	0.0024	0.0012
20	0.0217	0.0062	0.0011	0.0001	0
21	0.0974	0.0458	0.0132	0.0024	0.0003
22	0.1607	0.1244	0.0591	0.0174	0.0033
23	0.0974	0.1244	0.0974	0.0473	0.0147
24	0.0217	0.0458	0.0591	0.0473	0.0243
25	0.0018	0.0062	0.0132	0.0174	0.0147

	2.0	2.5	3.0	3.5	4.0
30	0.0132	0.0062	0.0018	0.0003	0
31	0.0591	0.0458	0.0217	0.0064	0.0012
32	0.0974	0.1244	0.0974	0.0473	0.0147
33	0.0591	0.1244	0.1607	0.1286	0.0661
34	0.0132	0.0458	0.0974	0.1286	0.1089
35	0.0011	0.0062	0.0217	0.0473	0.0661
40	0.0029	0.0023	0.0011	0.0003	0.0001
41	0.0132	0.0168	0.0132	0.0064	0.002
42	0.0217	0.0458	0.0591	0.0473	0.0243
43	0.0132	0.0458	0.0974	0.1286	0.1089
44	0.0029	0.0168	0.0591	0.1286	0.1796
45	0.0002	0.0023	0.0132	0.0473	0.1089
50	0.0002	0.0003	0.0002	0.0001	0
51	0.0011	0.0023	0.0029	0.0024	0.0012
52	0.0018	0.0062	0.0132	0.0174	0.0147
53	0.0011	0.0062	0.0217	0.0473	0.0661
54	0.0002	0.0023	0.0132	0.0473	0.1089
55	0	0.0003	0.0029	0.0174	0.0661

5

## 解析解 (T は連続とする)

$n = 1$  で  $X = x$  を観測した場合の最尤解は

$$L(t) = \exp(-0.5(x - t)^2)$$

を最大とする  $t$  は  $x$ .

$n = 2$  で  $X_1 = x_1 \& X_2 = x_2$  を観測した場合の最尤解は

$$L(t) = \exp(-0.5(x_1 - t)^2) \times \exp(-0.5(x_2 - t)^2)$$

を最大とする  $t$  は?

対数尤度は

$$\begin{aligned} \ln L(t) &= -0.5((x_1 - t)^2 + (x_2 - t)^2) \\ &= -0.5(2t^2 - 2(x_1 + x_2)t + x_1^2 + x_2^2) \\ &= -(t^2 - (x_1 + x_2)t) + t \text{ を含まない項} \end{aligned}$$

6

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。対数関数は単調増加関数であるため、対数尤度関数の最大値を与える母数はもとの尤度関数の最大値を与える。

母数が連続量の場合、最尤解を求めるためには対数尤度関数の母数  $\theta$  に関する微分を 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。

7

## 正規分布で $\sigma=1$ を仮定した場合の例

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

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

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

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

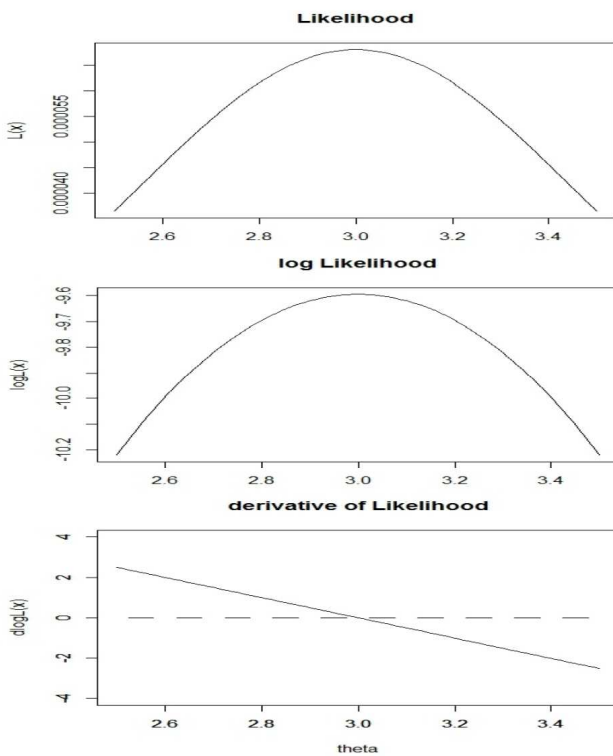
$$d \ln L(\theta) = \sum_{i=1}^n (y_i - \theta) = 0$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i$$

8

### 1.1.1 数値例（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

先ほどの表の転置で、データを  $Y$  として、  
 $Y_1 = 1 \& Y_2 = 2 \& Y_3 = 3 \& Y_4 = 4 \& Y_5 = 5$   
 の行のみ

## ベルヌーイ分布の例

一定の正規確率  $\theta$  を持つ独立な  $n$  回の試行

$$y_i = \{0|1\}, \quad i = 1, 2, \dots, n$$

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

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

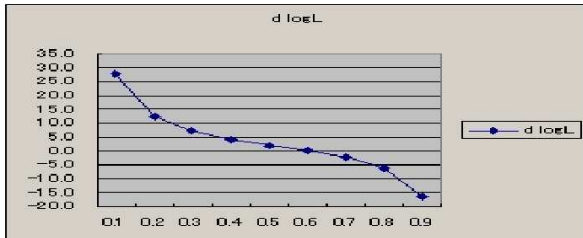
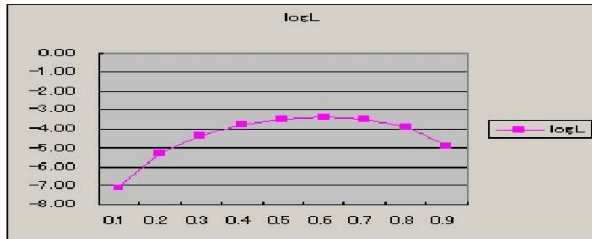
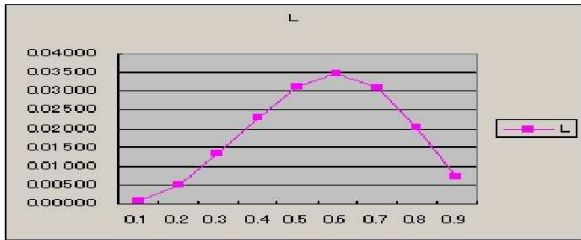
$$\ln L(\theta) = \sum_{i=1}^n (y_i \log \theta + (1 - y_i) \log(1 - \theta))$$

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

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n y_i$$

## 1.2.1 数値例（2項分布）

$y = \{0, 1, 1, 0, 1\}$  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

先ほどの表の転置で、データを  $Y$  として、 $Y_1 = 0 \& Y_2 = 1 \& Y_3 = 1 \& Y_4 = 0 \& Y_5 = 1$  の行のみ

11

## 最尤解の性質

母数  $\theta$  を

$$\tau = g(\theta)$$

と変換し新たな母数  $\tau$  を作ったとすると、対数尤度の最大値を与える  $\theta$  の値  $\hat{\theta}$  は対数尤度の最大値を与える  $\tau$  の値  $\hat{\tau}$  を

$$\hat{\theta} = g^{-1}(\hat{\tau})$$

と変換したものとなる。

なぜなら

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{\partial L(g^{-1}(\tau))}{\partial \tau} \frac{\partial \tau}{\partial \theta}$$

であるが、

$$\frac{\partial L(g^{-1}(\tau))}{\partial \tau} \Big|_{\tau=\hat{\tau}} = 0$$

だから。

12

## 正規分布で $\mu$ も $\sigma$ も未知の場合の例

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

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

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

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

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

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

$$\tau = 1/\sigma^2$$

13

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

$$\frac{\partial \ln L}{\partial \sigma} = ???$$

$$\frac{\partial \ln L}{\partial \sigma^2} = ???$$

$$\frac{\partial \ln L}{\partial \tau} = -\frac{n}{\tau} + \sum_{i=1}^n (y_i - \mu)^2$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n y_i$$

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

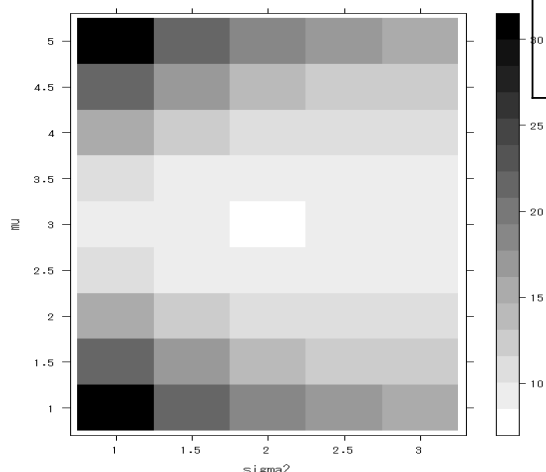
$$\hat{\sigma}^2 = \frac{1}{\hat{\tau}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\mu})^2$$

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

14

# 数値例

$$n \ln \sigma^2 + \sum_{i=1}^n (y_i - \mu)^2 / \sigma^2$$



	1.0	1.5	2.0	2.5	3.0
1.0	30.00	22.03	18.47	16.58	15.49
1.5	21.25	16.19	14.09	13.08	12.58
2.0	15.00	12.03	10.97	10.58	10.49
2.5	11.25	9.53	9.09	9.08	9.24
3.0	10.00	8.69	8.47	8.58	8.83
3.5	11.25	9.53	9.09	9.08	9.24
4.0	15.00	12.03	10.97	10.58	10.49
4.5	21.25	16.19	14.09	13.08	12.58
5.0	30.00	22.03	18.47	16.58	15.49

15

# 回帰分析

データ

$(x_i, y_i), i = 1, 2, \dots, n$  ただし  $y_i$  は連続量

モデル

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

または

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

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

尤度関数

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

対数尤度関数

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

16



## 対数尤度関数

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

これを最大にするように  $\boldsymbol{\xi}$  を求める。すなわち、連立方程式

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

を解く。答えは

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

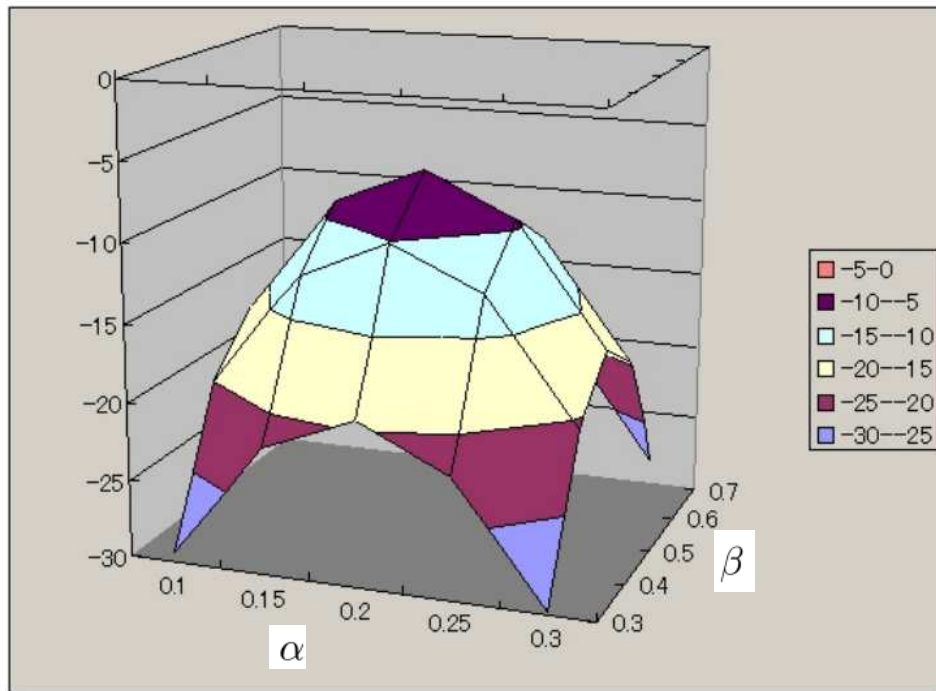
## 数値例

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

Data1			
Obs	X	y	yhat
1	-3.0	0.000	-0.200
2	-2.6	0.000	-0.080
3	-2.3	0.000	0.010
4	-1.9	0.000	0.130
5	-1.5	0.200	0.250
6	-1.1	0.166	0.370
7	-0.8	0.285	0.460
8	-0.4	0.125	0.580
9	0.0	0.777	0.700
10	0.4	0.875	0.820
11	0.8	0.714	0.940
12	1.1	1.000	1.030
13	1.5	0.800	1.150
14	1.9	1.000	1.270
15	2.3	0.666	1.390
16	2.6	1.000	1.480
17	3.0	1.000	1.600

最尤解の探索		
alpha	beta	LL
0.3	0.1	-29.316
0.3	0.15	-21.841
0.3	0.2	-19.438
0.3	0.25	-22.105
0.3	0.3	-29.843
0.4	0.1	-19.984
0.4	0.15	-12.509
0.4	0.2	-10.106
0.4	0.25	-12.773
0.4	0.3	-20.511
0.5	0.1	-16.614
0.5	0.15	-9.139
0.5	0.2	-6.735 *
0.5	0.25	-9.402
0.5	0.3	-17.141
0.6	0.1	-19.205
0.6	0.15	-11.730
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



19

## ロジスティク回帰

データ

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

これを  $x_i$  の異なる値ごとに度数分布の形にしたもの

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

モデル

確率は  $(0, 1)$  の区間に入らなければならないので、

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

とする。これは

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

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

$$\begin{aligned} \mathcal{E}(Y_i|x_i) &= P(x_i|\alpha, \beta) = \pi_i \\ \text{var}(Y_i|x_i) &= \pi_i(1 - \pi_i) \end{aligned}$$

となる。また、 $(f_i, r_i)$  を試行数と度数として、

$$\begin{aligned} \mathcal{E}(R_i|x_i) &= f_i \times P(x_i|\alpha, \beta) = f_i \pi_i \\ \text{var}(R_i|x_i) &= f_i \pi_i(1 - \pi_i) \end{aligned}$$

20

## 尤度関数

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

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

## 対数尤度関数

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

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

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

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

を解く。答えは解析的には解けない。

21

## 数値例

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

### 最尤解

$\alpha$  1.4642 (0.861)  
 $\beta$  0.2127 (-0.14)

22

