第 V 部　Logistic Model

2　最尤推定

ポイント

- 尤度
- 対数尤度関数
- 完全分離と準完全分離

2.1　はじめに

- 神様は平等です。
- 意地悪も依怙贔屓（えこひいき）もしません。
- 観測結果は普通のことなのです。
- 観測結果を『最（もっと）も』得られやすい値が「真の値」として『尤（もっと）も』らしいのです。
- 最も尤もらしい値を選ぶのが『最尤法（さいゆうほう）』です。

2.2　尤度

- y を確率変数とする。
- y をまだ観測していない状況で、観測されうる事象を想定し、それぞれに確率の値を対応させる。
- この時点では事象を特定できていないので『確率』という言葉は相応しくない。
- その後、施行を行い、事象を観測する。
- この時点で事象が確定し、それによって値が特定される。
- このような値を『尤度（ゆうど）』という。

歪んだコインを使った説明

- 歪んだコインを投げ 表 がでたら $y = 1$、裏が出たら $y = 0$ とする。
- y はコインの表裏によって決まる確率変数であって、1 または 0 の値しかとらない。
- $y = 1$ のときの確率を p とし、$y = 0$ の時の確率を $1 - p$ とする。
- コインを投げると表か裏が出る。
- コインを投げた後の結果によって p か $1 - p$ が決まる。
- 観測後に確定した y によって決まる値が『尤度』である。

2.2.1　y がダミー変数のときの尤度

- y を 1 または 0 しかとらない確率変数であるとする。
- $y = 1$ のときの確率 $\Pr(y = 1)$ を p、$y = 0$ の時の確率 $\Pr(y = 0)$ を $1 - p$ とする。
- y の尤度を

$$l_i = {p_i}^{y_i} \left(1 - p_i\right)^{1 - y_i} \tag{1}$$

と定義する。

2.2.2 y の出現値による尤度の違い

$y = 1$ の場合

$$l_i = {p_i}^1 \left(1 - p_i\right)^{1-1} = p_i \tag{2}$$

$y = 0$ の場合

$$l_i = {p_i}^0 \left(1 - p_i\right)^{1-0} = 1 - p_i \tag{3}$$

2.2.3 確率と尤度の違い

- 確率は、事象ごとに定められた値
- 尤度は、観測によって確定する事象に対応する値

2.3 尤度関数

- n 個の尤度を、それぞれ独立な事象の確率であると仮定する。
- この考えに立ち、観測値が得られる同時確率を求める。
- 観測結果は独立な事象なので同時確率は尤度の総乗である。
- この尤度 (l_i) の総乗を『尤度関数』とよび L であらわす。
- 尤度 (l_i) は p の関数なので尤度関数 (L) も p の関数である。

$$L = l_1 \times l_2 \times l_3 \times \cdots \times l_n = \prod_{i=1}^{n} l_i \tag{4}$$

2.4　最尤推定

- 観測結果が出現したのは、確率がそもそも観測結果を『最(もっと)も得られやすい値』であったからと仮定する。
- この考えに従うと尤度関数が最大となる尤度 (l_i) が観測値を得るために最(もっと)も『尤(もっと)もらしい』のである。
- 尤度関数を最大化するように尤度 (l_i) のパラメータ (p) を推定する推定法を『最尤推定(さいゆうすいてい)』または『最尤法(さいゆうほう)』という。
- 尤度 (l_i) はパラメータ (p) の関数なので尤度関数 (L) を最大化するようにパラメータ (p) を推定する。

2.4.1 歪んだコインを使った説明

歪んだコインを投げ表が出た場合 $y=1$、裏が出た場合 $y=0$ とする。$y=1$ のときの確率を p とすると $y=0$ の確率は $1-p$ である。

である。このコインを 8 回投げて出た結果をまとめたものが 表**??**である。

表 1 コインの表裏と尤度

i	観測値	y_i	l_i
1	表	1	p
2	裏	0	$1-p$
3	裏	0	$1-p$
4	裏	0	$1-p$
5	裏	0	$1-p$
6	裏	0	$1-p$
7	表	1	p
8	表	1	p

2.4.2 数値例に基づく尤度関数 L の定式化

- それぞれの尤度を独立な事象の確率とみなした同時確率が尤度関数 (L)
- 観測結果はそれぞれ独立な事象なので同時確率は尤度の総乗。

したがって尤度関数 (L) は

$$L = \prod_{i=1}^{8} l_i = l_1 \times l_2 \times l_3 \times l_4 \times l_5 \times l_6 \times l_7 \times l_8 \tag{5}$$

$$= p(1-p)(1-p)(1-p)(1-p)(1-p)pp \tag{6}$$

$$= p^3(1-p)^5 \tag{7}$$

この尤度関数 (L) を最大化するように p を求める推定が最尤推定である。

2.4.3 尤度関数を最大化するためのアイデア

- 尤度の範囲は $(0, 1)$ なので、その総乗は極めて微小な値 (限りなく 0 に近い値) になる。
- 微小な値の最大値を与える p を求めることは困難である。
- ネイピア数 e を底とする対数関数は狭義単調増加関数なので、対数尤度関数 $(\ln L)$ を最大化する p は尤度関数 (L) を最大化する。
- この性質を利用し尤度関数 (L) を対数変換した対数尤度関数 $(\ln L)$ を最大化する。
- 対数尤度関数の最大化は逐次探索的に行われる。

問題 V–2 – 1

以下の尤度関数 L を対数変換しなさい。ただし、$0 < p < 1$ とする。

$$L = p^3(1-p)^5$$

解例 V−2 − 1

仮定により、

$$0 < p < 1$$

なので

$$0 < 1 - p < 1$$

である。したがって

$$0 < L$$

$$\begin{aligned}\ln L &= \ln\left(p^3(1-p)^5\right)\\ &= \ln\left(p^3\right) + \ln\left((1-p)^5\right)\\ &= 3\ln p + 5\ln(1-p)\end{aligned}$$

表 2 $(0.0, 1.0)$ **の範囲の** L **と** $\ln L$

p	L	$\ln L$
0.2	0.00262144	−5.94403149
0.3	0.00453789	−5.39529313
0.4	0.00497664	−5.30300031
0.5	0.00390625	−5.54517744
0.6	0.00221184	−6.11393053
0.7	0.00083349	−7.08988885
0.8	0.00016384	−8.71662022

表 3 $(0.3, 0.5)$ **の範囲の** L **と** $\ln L$

p	L	$\ln L$
0.34	0.00492217	−5.31400620
0.36	0.00500965	−5.29638926
0.38	0.00502700	−5.29293108
0.40	0.00497664	−5.30300031
0.42	0.00486282	−5.32613758
0.44	0.00469135	−5.36203413
0.46	0.00446933	−5.41051707

図 1 尤度関数 L

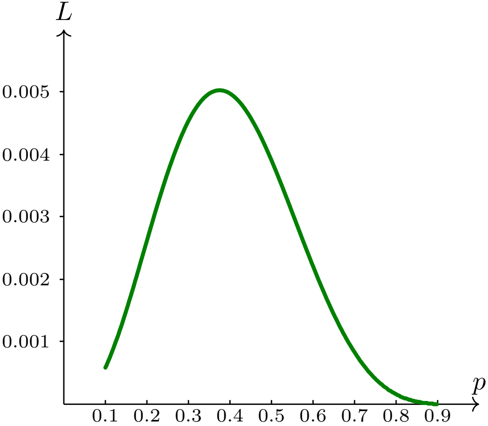
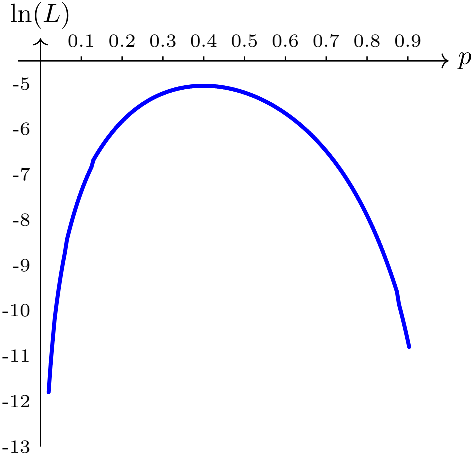


図 2 対数尤度関数 $\ln(L)$



問題 V−2 − 2

以下の尤度関数 L を対数変換しなさい。ただし、$0 < p < 1$ とする。

$$L = \prod_{i=1}^{3} {p_i}^{y_i} \left(1 - p_i\right)^{1-y_i}$$

解例 V–2 – 2

尤度関数を総乗記号を使わずにあらわすと。

$$
\begin{aligned}
L &= \prod_{i=1}^{3} {p_i}^{y_i} \left(1 - p_i\right)^{1-y_i} \\
&= {p_1}^{y_1} \left(1 - p_1\right)^{1-y_1} \times {p_2}^{y_2} \left(1 - p_2\right)^{1-y_2} \times {p_3}^{y_3} \left(1 - p_3\right)^{1-y_3}
\end{aligned}
$$

対数変換すると

$$
\ln L = \ln \left({p_1}^{y_1} \left(1 - p_1\right)^{1-y_1} {p_2}^{y_2} \left(1 - p_2\right)^{1-y_2} {p_3}^{y_3} \left(1 - p_3\right)^{1-y_3} \right)
$$

真数の因数をそれぞれ真数とする対数の和であらわし

$$
\begin{aligned}
= \quad & \ln {p_1}^{y_1} + \ln \left(1 - p_1\right)^{1-y_1} \\
& + \ln {p_2}^{y_2} + \ln \left(1 - p_2\right)^{1-y_2} \\
& + \ln {p_3}^{y_3} + \ln \left(1 - p_3\right)^{1-y_3}
\end{aligned}
$$

真数の冪数を係数に掃きだし

$$
\begin{aligned}
= \quad & y_1 \ln p_1 + (1 - y_1) \ln (1 - p_1) \\
& + y_2 \ln p_2 + (1 - y_2) \ln (1 - p_2) \\
& + y_3 \ln p_3 + (1 - y_3) \ln (1 - p_3)
\end{aligned}
$$

総和記号でまとめ

$$
\ln L = \sum_{i=1}^{3} \left(y_i \ln p_i + (1 - y_i) \ln (1 - p_i)\right)
$$

問題 V–2 – 3

$$p_i = \frac{1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)}$$

とする。$1 - p_i$ を求めなさい。

解例 V−2 − 3

$$
\begin{aligned}
1 - p_i &= 1 - \frac{1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)} \\
&= \frac{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right) - 1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)} \\
&= \frac{\exp\left(-\left(\alpha + \beta x_i\right)\right)}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)}
\end{aligned}
$$

$\dfrac{\exp(\alpha + \beta x_i)}{\exp(\alpha + \beta x_i)} = 1$ を掛け

$$
= \frac{1}{1 + \exp\left(\alpha + \beta x_i\right)}
$$

2.5　ロジスティックモデルによる尤度関数と対数尤度関数

$P(y=1)$ にロジスティックモデル

$$p_i = \frac{1}{1+\exp\left(-\left(\alpha+\beta x_i\right)\right)} \tag{8}$$

を仮定する。すると $P(y=0)$ は

$$1-p_i = \frac{1}{1+\exp\left(\alpha+\beta x_i\right)} \tag{9}$$

である。

すると尤度は

$$l_i = \left(\frac{1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)} \right)^{y_i} \left(\frac{1}{1 + \exp\left(\alpha + \beta x_i\right)} \right)^{1-y_i} \tag{10}$$

であり、尤度関数は

$$L = \prod_{i=1}^{n} \left(\frac{1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)} \right)^{y_i} \left(\frac{1}{1 + \exp\left(\alpha + \beta x_i\right)} \right)^{1-y_i} \tag{11}$$

である。

(**??**) を対数変換すると対数尤度関数

$$\ln L = \sum_{i=1}^{n} \left(y_i \ln \left(\frac{1}{1 + \exp\left(-\left(\alpha + \beta x_i\right)\right)} \right) + (1 - y_i) \ln \left(\frac{1}{1 + \exp\left(\alpha + \beta x_i\right)} \right) \right) \tag{12}$$

を得る。

ロジスティックモデルの対数尤度関数 ($\ln L$) は α と β の関数である。従って、最尤推定は $\ln L$ を最大化するように α と β を推定する。

2.6 数値例

ある製品を作成する為には加熱処理が必要であるが、加熱処理の時間 (x) によって不良品 $(y = 1)$ が発生しやすくなることが想定されている。

そこで加熱時間 (x) を変えながら 100 個ずつ作成し、不良品 $(y = 1)$ か否か $(y = 0)$ を確認した。

観測結果が表??である。

表 4　加熱時間と不良品と正常品

x	$y = 1$	$y = 0$
10	2	98
11	8	92
12	27	73
13	62	38
14	88	12
15	97	3
16	99	1

問題 V–2 – 4

$x = 13$ のとき $y = 1$ は 62 回観測され、$y = 0$ は 38 回観測されたとする。
p_i にロジスティックモデルを仮定し、尤度関数, 対数尤度関数を示しなさい。

$$p_i = \frac{1}{1 + \exp\big(-(\alpha + \beta x_i)\big)}$$

$$1 - p_i = \frac{1}{1 + \exp\big(\alpha + \beta x_i\big)}$$

解例 V−2 − 4

$$l_i = \left(\frac{1}{1+\exp(-(\alpha+(13)\beta))}\right)^{y_i}\left(\frac{1}{1+\exp(\alpha+(13)\beta)}\right)^{1-y_i}$$

$$L = \left(\frac{1}{1+\exp\left(-\left(\alpha+(13)\beta\right)\right)}\right)^{62}\left(\frac{1}{1+\exp\left(\alpha+(13)\beta\right)}\right)^{38}$$

$$\ln L = 62\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(13)\beta\right)\right)}\right)+38\ln\left(\frac{1}{1+\exp\left(\alpha+(13)\beta\right)}\right)$$

2.6.1 x ごとの対数尤度の和

$x=10$ のときの対数尤度の和を $\ln L_{10}$ とあらわす。$x=11$ 以降も同様にあらわす。

$$\ln L_{10} = 2\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(10)\beta\right)\right)}\right)+98\ln\left(\frac{1}{1+\exp\left(\alpha+(10)\beta\right)}\right) \tag{13}$$

$$\ln L_{11} = 8\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(11)\beta\right)\right)}\right)+92\ln\left(\frac{1}{1+\exp\left(\alpha+(11)\beta\right)}\right) \tag{14}$$

$$\ln L_{12} = 27\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(12)\beta\right)\right)}\right)+73\ln\left(\frac{1}{1+\exp\left(\alpha+(12)\beta\right)}\right) \tag{15}$$

$$\ln L_{13} = 62\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(13)\beta\right)\right)}\right)+38\ln\left(\frac{1}{1+\exp\left(\alpha+(13)\beta\right)}\right) \tag{16}$$

$$\ln L_{14} = 88\ln\left(\frac{1}{1+\exp\left(-\left(\alpha+(14)\beta\right)\right)}\right)+12\ln\left(\frac{1}{1+\exp\left(\alpha+(14)\beta\right)}\right) \tag{17}$$

$$\ln L_{15} = 97 \ln \left(\frac{1}{1 + \exp\left(-\left(\alpha + (15)\beta\right)\right)} \right) + \ 3 \ln \left(\frac{1}{1 + \exp\left(\alpha + (15)\beta\right)} \right) \tag{18}$$

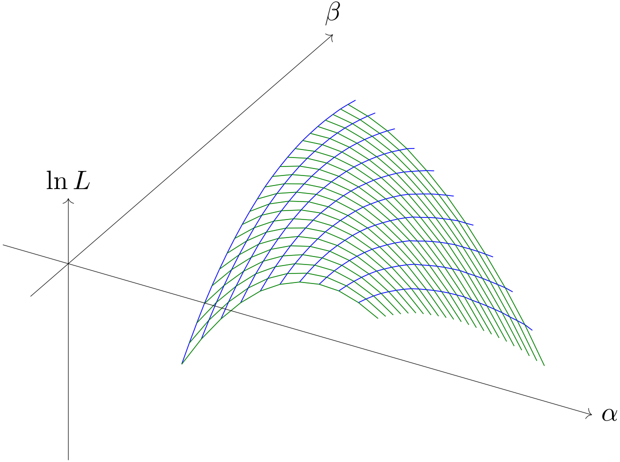
$$\ln L_{16} = 99 \ln \left(\frac{1}{1 + \exp\left(-\left(\alpha + (16)\beta\right)\right)} \right) + \ 1 \ln \left(\frac{1}{1 + \exp\left(\alpha + (16)\beta\right)} \right) \tag{19}$$

これらの「x ごとの対数尤度の和」の総和が対数尤度関数 $(\ln L)$

$$\ln L = \ln L_{10} + \ln L_{11} + \ln L_{12} + \ln L_{13} + \ln L_{14} + \ln L_{15} + \ln L_{16} \tag{20}$$

である。

図 3　$\ln L$ の形状



問題 V–2 – 5

表の値を使い最尤法でパラメータ α, β を推定しなさい。
ただし

$$p_i = \frac{1}{1 + \exp(-(\alpha + \beta x_i))}$$
$$1 - p_i = \frac{1}{1 + \exp(\alpha + \beta x_i)}$$

とする。

x	$y = 1$	$y = 0$
10	2	98
11	8	92
12	27	73
13	62	38
14	88	12
15	97	3
16	99	1

解例 V−2 − 5

$\hat{\alpha}$	$-18.599\,425\,486$
$\hat{\beta}$	$1.468\,415\,317$
$\max(\ln L)$	$-218.233\,482\,714$

問題 V–2 – 6

表の値を使い最尤法でパラメータ α, β を推定しなさい。ただし

$$p_i = \frac{1}{1 + \exp(-(\alpha + \beta x_i))}$$
$$1 - p_i = \frac{1}{1 + \exp(\alpha + \beta x_i)}$$

とする。

x	$y = 1$	$y = 0$
10	0	100
11	8	92
12	27	73
13	62	38
14	88	12
15	97	3
16	99	1

解例 V−2 − 6

$\hat{\alpha}$	$-19.435\,219\,414$
$\hat{\beta}$	$1.531\,870\,619$
$\max(\ln L)$	$-210.205\,064\,036$

問題 V–2 – 7

表の値を使い最尤法でパラメータ α, β を推定しなさい。ただし

$$p_i = \frac{1}{1+\exp(-(\alpha+\beta x_i))}$$
$$1-p_i = \frac{1}{1+\exp(\alpha+\beta x_i)}$$

とする。

x	$y=1$	$y=0$
10	2	98
11	8	92
12	27	73
13	62	38
14	88	12
15	97	3
16	100	0

解例 V−2 − 7

$\hat{\alpha}$	$-19.070\,296\,044$
$\hat{\beta}$	$1.506\,735\,578$
$\max(\ln L)$	$-213.267\,756\,426$

2.7　最尤推定ができない場合

以下の場合、パラメータの推定はできない。

1. $\forall y_i = 1$　あるいは　$\forall y_i = 0$
2. $A = \{x_i | y_i = 0\}$

   $B = \{x_i | y_i = 1\}$

   において

   (a) $\max(A) < \min(B)$　あるいは　$\max(B) < \min(A)$

   (b) $\max(A) = \min(B)$　あるいは　$\max(B) = \min(A)$

2.7.1 $\forall y_i = 1$ の場合

$y_i = 1$ のときの尤度 (l_i) は

$$l_i = \frac{1}{1+\exp(-(\alpha+\beta x_i))} \tag{21}$$
$$= \frac{1}{1+\exp\left(-\beta\left(x_i-\left(-\frac{\alpha}{\beta}\right)\right)\right)} \tag{22}$$

である。

(??) において、$\beta > 0$ とする。

$$-\frac{\alpha}{\beta} < \min(x) \tag{23}$$

となるように α, β をとると、ロジスティックモデルの「対称の中心」の x 軸上の座標は $\min(x)$ よりも (–) 方向にある。α はいくらでも大きくとることができるから

$$\alpha \to +\infty \implies l_i \to 1 \tag{24}$$

である。

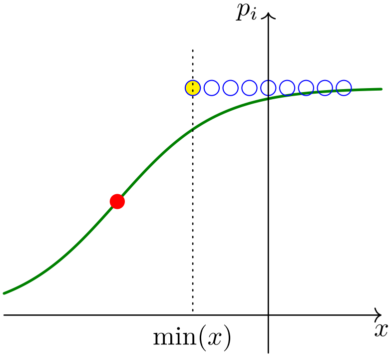
また β においても「対称の中心」が $\min(x)$ の (–) 方向にある状況であれば、

$$\beta \to +\infty \implies l_i \to 1 \tag{25}$$

である。

全ての l_i は $\alpha,\ \beta$ に関して狭義単調増加関数であり、その総乗である尤度関数 (L) は最大値を持たない。したがってパラメータを推定することができない。この状況を「パラメータが発散する」という。

図 4　$\forall \boldsymbol{y_i} = \mathbf{1},\ \boldsymbol{\beta} > \mathbf{0}$



次に $\beta < 0$ とする。

$$-\frac{\alpha}{\beta} > \max(x) \tag{26}$$

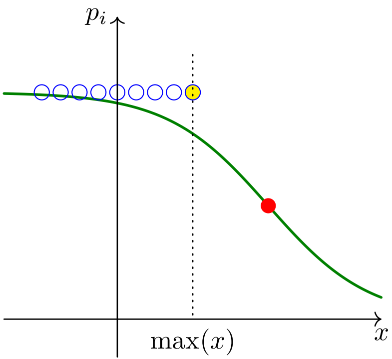
となるように α, β をとると、「対称の中心」の x 軸上の座標は $\max(x)$ よりも $(+)$ 方向にある。従って

$$\alpha \to -\infty \implies l_i \to 1 \tag{27}$$

$$\beta \to -\infty \implies l_i \to 1 \tag{28}$$

なので、尤度関数 (L) は最大値を持たない。

図 5 $\forall y_i = 1$, $\beta < 0$



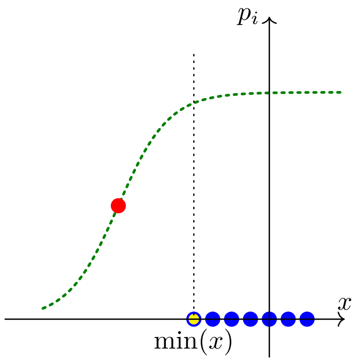
2.7.2 $\forall y_i = 0$ の場合

$y_i = 0$ のときの尤度 l_i は

$$l_i = \frac{1}{1 + \exp(\alpha + \beta x_i)} \tag{29}$$

である。尤度 (l_i) は α, β に関して傾向は異なるがともに狭義単調関数なので最大値を持たない。

図 6 $\forall y_i = 0$, $\beta < 0$



2.7.3　$x = \gamma$ で分けられる場合

$\max(A) < \min(B)$ の場合

$$A = \{x_i \mid y_i = 0\} \tag{30}$$

$$B = \{x_i \mid y_i = 1\} \tag{31}$$

において

$$\max(A) < \min(B) \tag{32}$$

とする。すると $\beta > 0$ として、

$$\max(A) < -\frac{\alpha}{\beta} < \min(B) \tag{33}$$

となるように α, β をとることができる。この点を γ とする。すると

$$\gamma = -\frac{\alpha}{\beta} \tag{34}$$

$$\alpha = -\gamma\beta \tag{35}$$

である。

　ここで、x_i の添え字を振りなおし、

$$x_j \in A \tag{36}$$

$$x_k \in B \tag{37}$$

とする。

　このとき「対称の中心」は γ にある。ここで $\beta \to +\infty$ とすると、

$$\beta \to +\infty \implies 1 - p_j \to 1, \quad p_k \to 1 \tag{38}$$

である。なので、パラメータは発散する。

$\max(B) < \min(A)$ の場合も同様である。

　このことを『完全分離』という。

$\max(A) = \min(B)$ の場合

$$\max(A) = \min(B) \tag{39}$$

とする。$\beta > 0$ として

$$\max(A) = \min(B) = -\frac{\alpha}{\beta} \tag{40}$$

を満たす α, β を γ とすると、同様の理由によりパラメータは発散する。

$\max(B) = \min(A)$ の場合も同様である。

　このことを『準完全分離』という。

図 7　完全分離

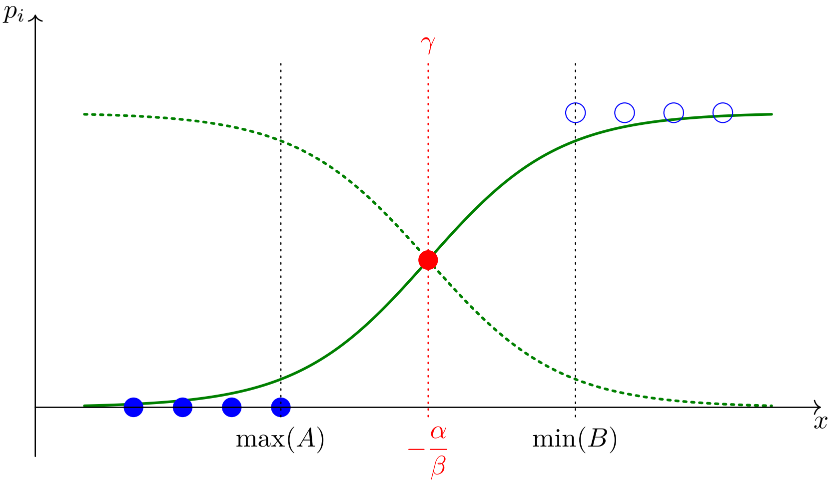
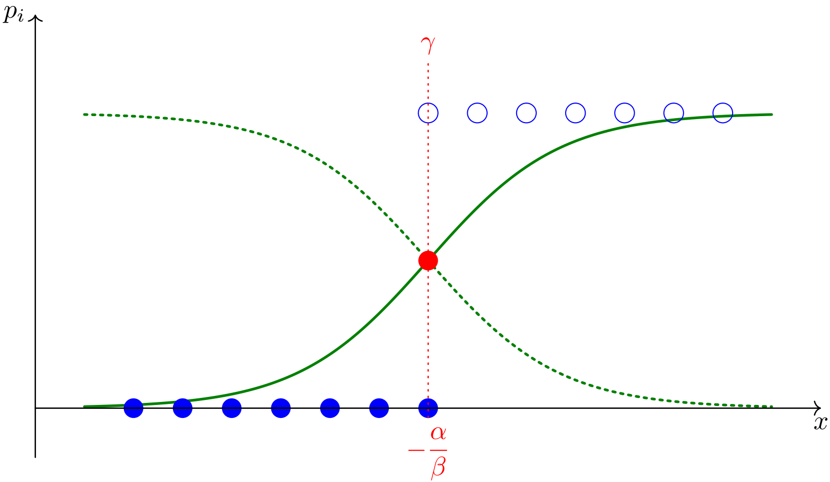


図 8 準完全分離



2.8　まとめ

- 尤度とは、観測する前に設定している値。
- 尤度関数 (L) とは尤度を独立な事象の確率であると仮定した同時確率。
- 尤度関数 (L) は微小な値になるので対数尤度関数 ($\ln L$) で対応。
- 尤度関数 (L) を対数変換したものが対数尤度関数 ($\ln L$)。
- 対数尤度関数 ($\ln L$) の最大のとき尤度関数 (L) も最大。
- 尤度関数を最大化するのが最尤法。
- 「完全分離」「準完全分離」のときは推定できない。