連鎖解析

東・荒井・鈴木・瀬沼
竹内・鳩谷・渡部

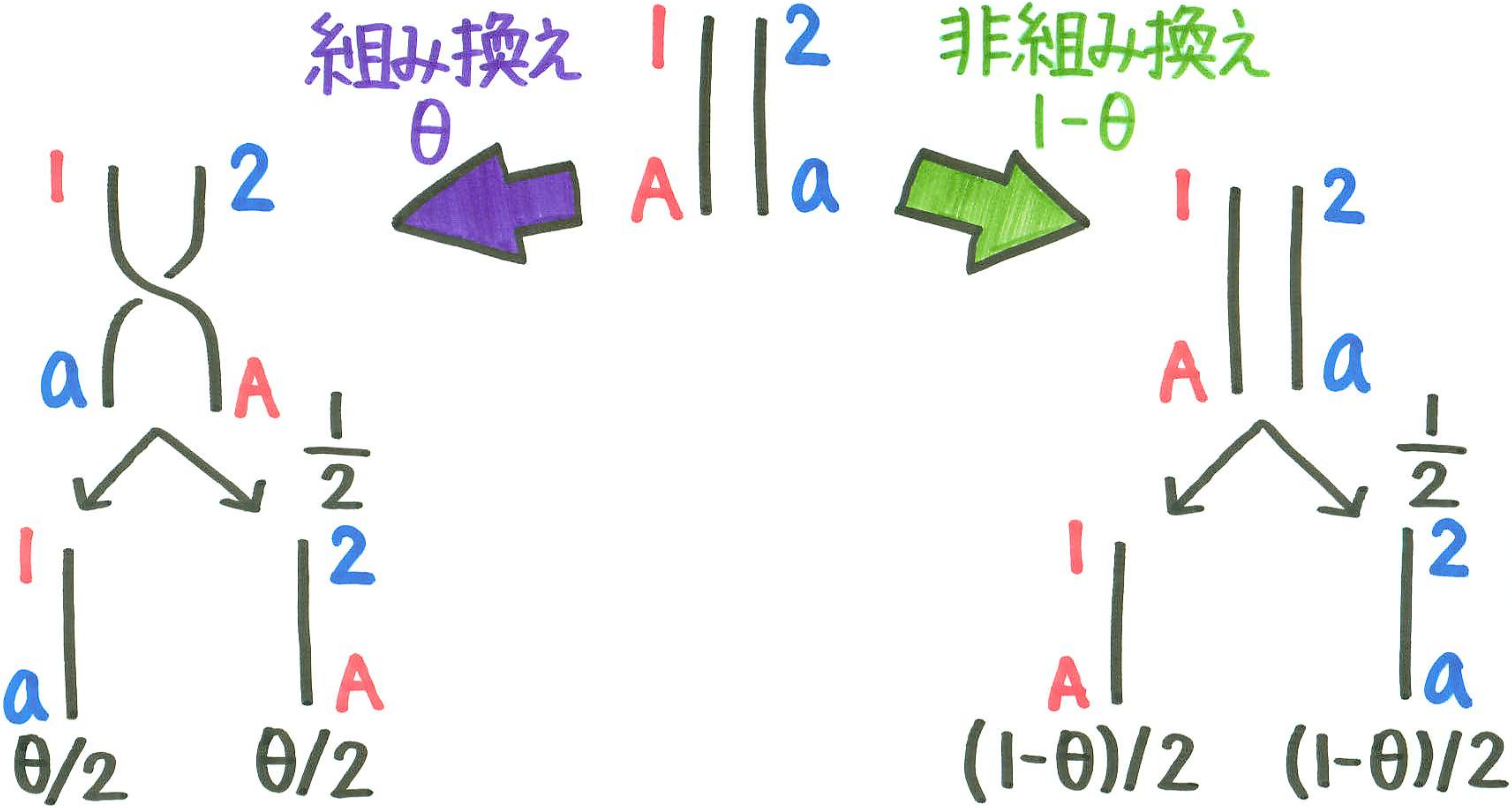
発表内容

連鎖解析 < パラメトリック連鎖解析 / ノンパラメトリック連鎖解析

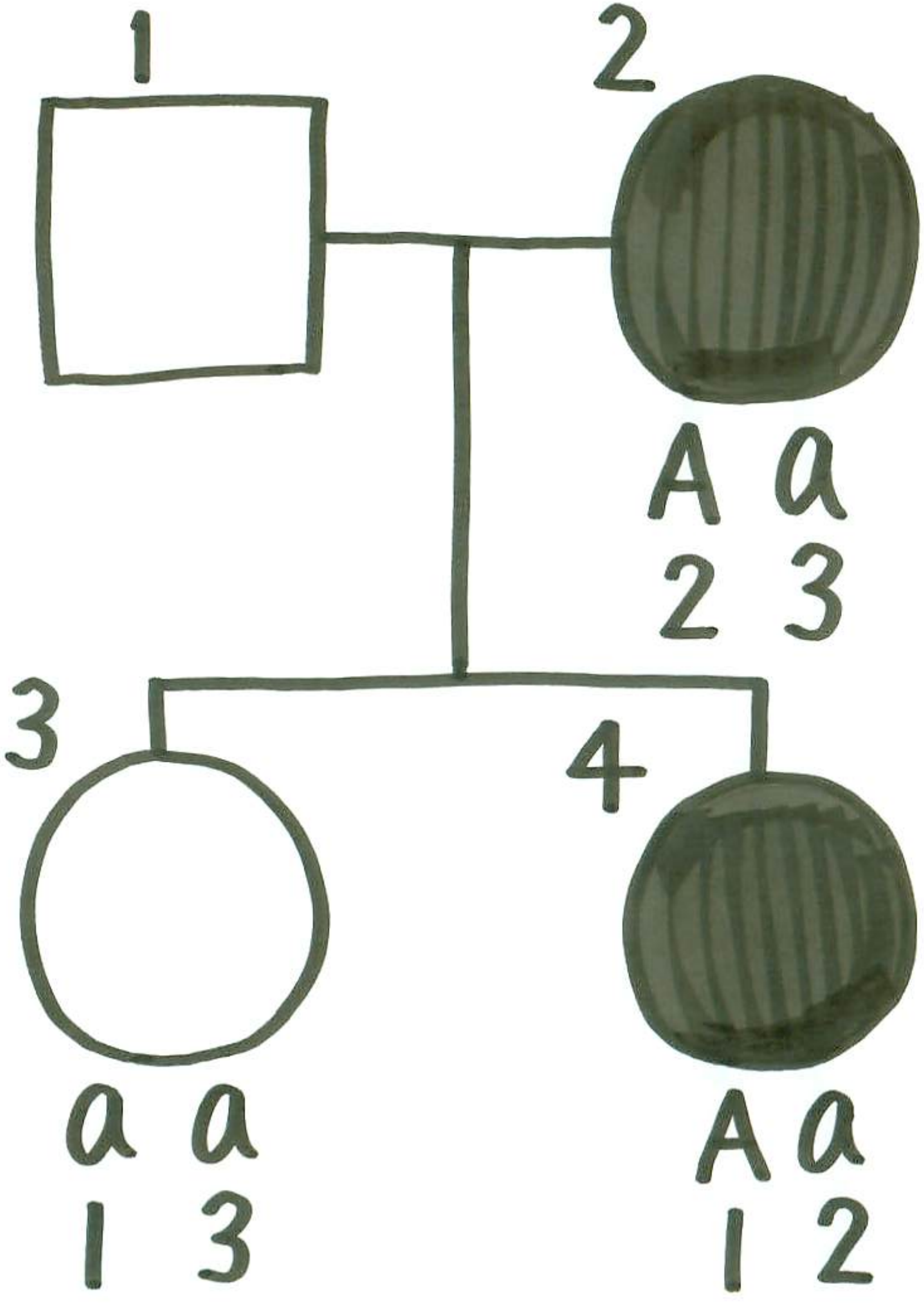
定義：Hardy-Weinberg均衡
マーカー対立遺伝子の確率
家系尤度
lod score関数

一部マーカー情報が欠損した家系図を用いた一例

組み換え・非組み換えと組み換え比



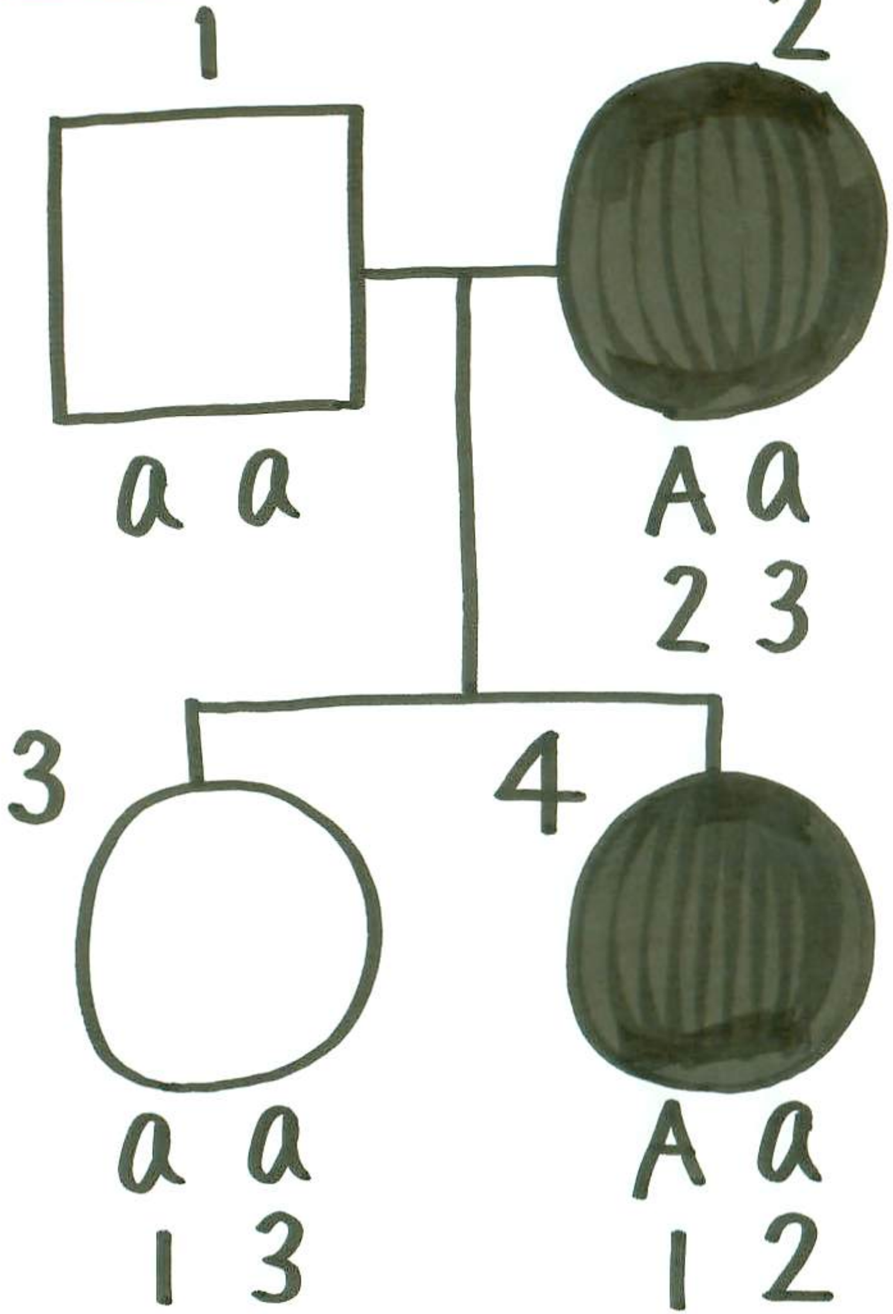
一部マーカー情報が欠損した家系



遺伝子モデルの仮定

- 病気が優性遺伝
- フェノコピーが存在しない
- 病気の対立遺伝子は頻度が低い

フェーズ①



母(2):

2 ‖ 3
A ‖ a
(2A|3a)

2 ‖ 3
a ‖ A
(2a|3A)

娘(4):

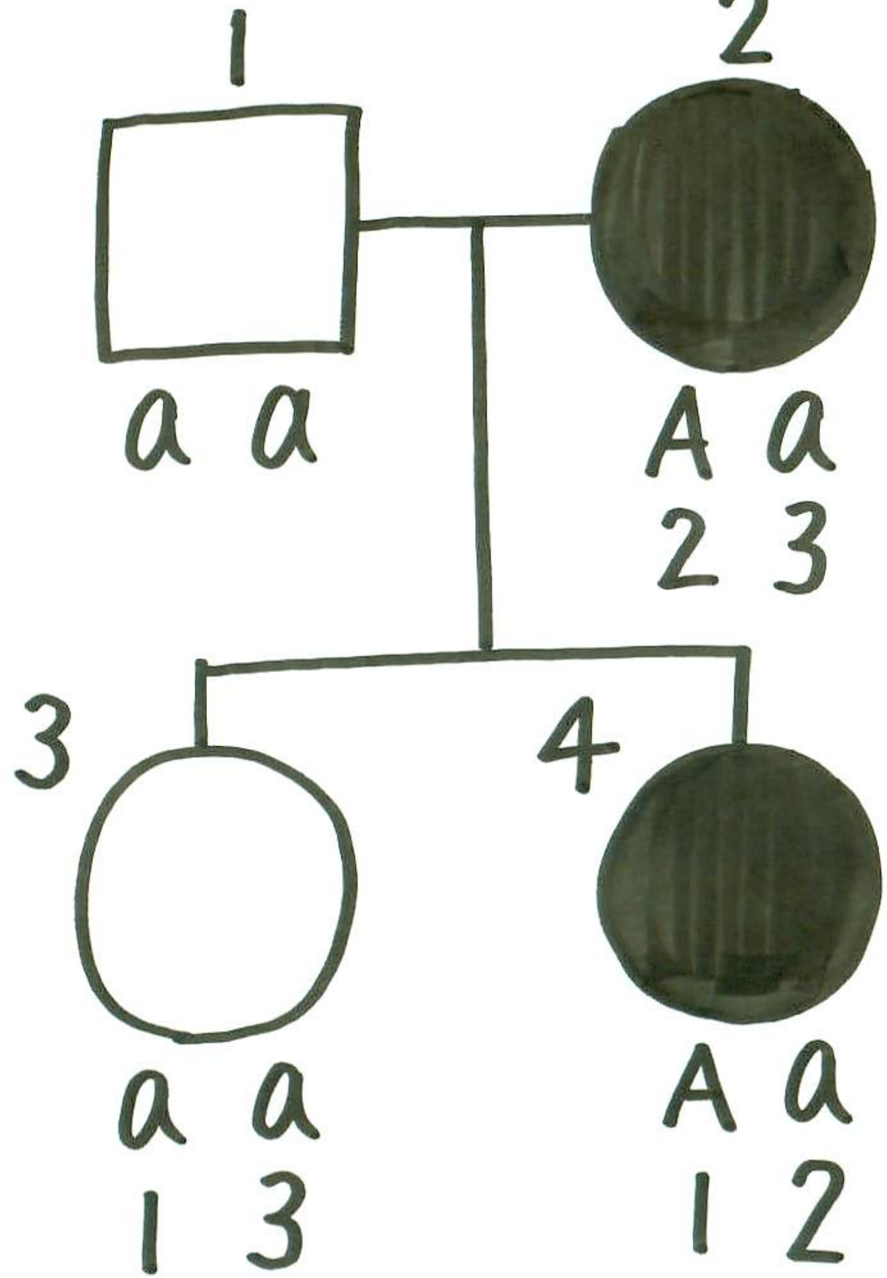
~~1 ‖ 2~~
~~A ‖ a~~

1 ‖ 2
a ‖ A
(1a|2A)

↓

父(1)は少なくとも一つ 1a をもつ

フェーズ②



母(2): (2a|3A)
(2A|3a)

娘(4): (1a|2A)

娘(3): (1a|3a)

父(1): (1a|1a)
(1a|2a)
(1a|3a)
(1a|wa)

G_1: (1a|1a, 2A|3a, 1a|3a, 1a|2A)

G_2: (1a|1a, 2a|3A, 1a|3a, 1a|2A)

G_3: (1a|2a, 2A|3a, 1a|3a, 1a|2A)

G_4: (1a|2a, 2a|3A, 1a|3a, 1a|2A)

G_5: (1a|3a, 2A|3a, 1a|3a, 1a|2A)

G_6: (1a|3a, 2a|3A, 1a|3a, 1a|2A)

G_7: (1a|wa, 2A|3a, 1a|3a, 1a|2A)

G_8: (1a|wa, 2a|3A, 1a|3a, 1a|2A)

6

Hardy-Weinberg均衡

$\begin{cases} A: \text{病気である対立遺伝子} \\ a: \text{病気でない対立遺伝子} \end{cases}$

$P(A)=p$, $P(a)=q$ のとき

ただし、$0 \leq p, q \leq 1$, $p+q=1$ とすると、

$$P((aa))=q^2$$
$$P((Aa))=2pq$$
$$P((AA))=p^2$$

を満たす。

マーカー番号の確率

$P(m=j)=p_j$ とする。

※今回の家系の場合

$j=1,2,3,w$

(w:家系図に表れた1~3以外の対立遺伝子)

ただし, $0 \leq p_1, p_2, p_3, p_w \leq 1$,

$p_1+p_2+p_3+p_w=1$ である。

家系尤度

F: founder , NF: nonfounder

g_i: 個体iの遺伝子型

g_{F_j}: 個体jから見た父の遺伝子型

g_{M_j}: 個体jから見た母の遺伝子型

y_i: 個体iの表現型

$$L(\theta)=\sum_{g}\prod_{i=1}^{n}P(y_i|g_i)\prod_{i\in F}P(g_i)\prod_{j\in NF}P(g_j|g_{F_j},g_{M_j})$$

※今回の家系の場合

$n=4$, $i=1,2$, $j=3,4$ である。

G_3の尤度

$$G_3 : (1a|2a,\ 2A|3a,\ 1a|3a,\ 1a|2A)$$

父(1) : $P(g_1) = 2p_1 p_2 \times q^2$

母(2) : $P(g_2) = 2p_2 p_3 \times 2pq$

娘(3) : $P(g_3 | g_1, g_2) = \dfrac{1}{2} \times \dfrac{(1-\theta)}{2}$

娘(4) : $P(g_4 | g_1, g_2) = \dfrac{1}{2} \times \dfrac{(1-\theta)}{2}$

G_3の尤度 : $\underbrace{1}_{\prod} \times \underbrace{2p_1 p_2 \times q^2 \times 2p_2 p_3 \times 2pq}_{=} \times \underbrace{\frac{1}{2} \times \frac{(1-\theta)}{2} \times \frac{1}{2} \times \frac{(1-\theta)}{2}}_{=}$

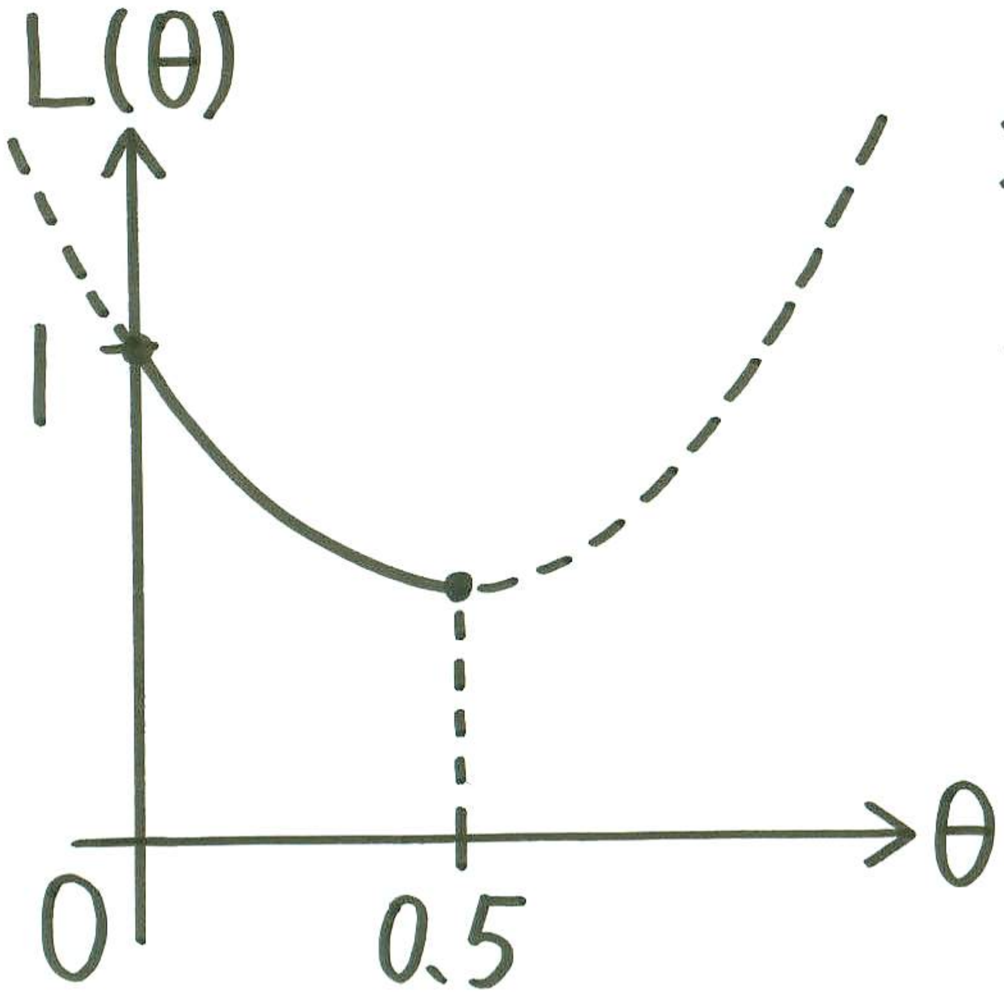
$$\prod_{i=1}^{n} P(y_i | g_i) \prod_{i \in F} P(g_i) \prod_{j \in NF} P(g_j | g_{F_j}, g_{M_j})$$

尤度関数

$$
\begin{aligned}
L(\theta) = 4p_1 q^2 \times 2p_2 p_3 \cdot 2pq \\
\times \Big\{ p_1 \cdot (1-\theta)^2 + p_1 \cdot \theta^2 + \frac{1}{2} p_2 (1-\theta)^2 \\
+ \frac{1}{2} p_2 \theta^2 + \frac{1}{2} p_3 (1-\theta)^2 + \frac{1}{2} p_3 \theta^2 \\
+ \frac{1}{2} p_w (1-\theta)^2 + \frac{1}{2} p_w \theta^2 \Big\} \\
\propto \theta^2 + (1-\theta)^2
\end{aligned}
$$

lod score 関数

$$Z(\theta) = \log_{10} \frac{L(\theta)}{L(0.5)} \quad (0 \leq \theta \leq 0.5)$$



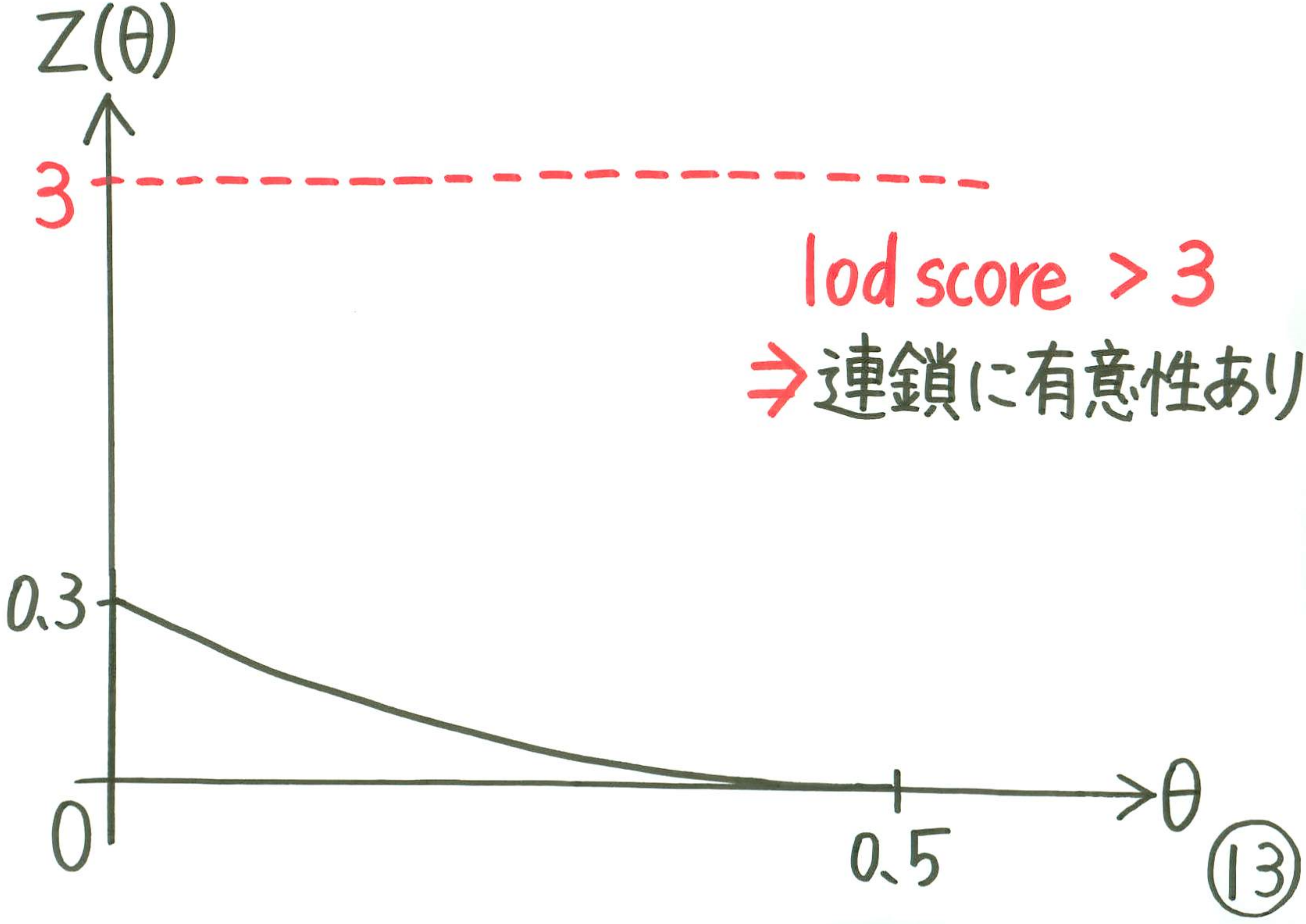
最大 lod score

$$Z(\theta=0) = \log_{10} \frac{L(0)}{L(0.5)}$$

$$= \log_{10} \frac{1}{0.5}$$

$$= \log_{10} 2$$

$$\fallingdotseq 0.3$$



Z(θ)
3
lod score > 3
⇒連鎖に有意性あり
0.3
0
0.5
θ
13