

摂動化によって プライバシーを保護した情報推薦方式

東海大学大学院 工学研究科 情報理工学専攻

OBDRM039 望月 安菜

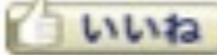
指導教員 菊池 浩明

情報推薦



耳をすませば [DVD]

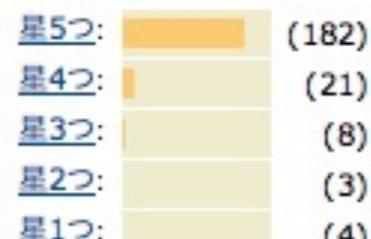
近藤喜文 (監督) | 形式: DVD

★★★★☆ (218件のカスタマーレビュー)  いいね (22)

カスタマーレビュー

耳をすませば [DVD]

218レビュー



おすすめ度

★★★★☆ (218件のカスタマーレビュー)

あなたの意見や感想を教えてください

[自分のレビューを作成する](#)

これにも注目

最近チェックした商品

同じテーマの商品



猫の恩返し/ギブリーズ
episode 2

森田宏幸, 池脇千鶴, 袴田吉彦, ...

VHS



猫の恩返し/ギブリーズ
episode2

森田宏幸, 百瀬義行, 宮崎駿

DVD

¥4,935 ¥ 3,663



崖の上のポニョ [DVD]

宮崎駿

DVD

¥4,935 ¥ 3,150



ゲド戦記 [DVD]

宮崎吾朗

DVD

¥4,935 ¥ 2,980



ハウルの動く城 [DVD]

宮崎駿

DVD

¥4,935 ¥ 3,700

- ◆ “摂動化によってプライバシーを保護した情報推薦方式”, DICOMO2011, 優秀論文賞受賞
- ◆ **“Slope One を用いた摂動化プライバシー保護情報推薦方式”, CSS2011**
- ◆ “アイテム依存の摂動化によるプライバシー保護情報推薦”, SCIS2012

摂動化プライバシー保護

	i_1	i_2	i_3
u_1	2	3	
u_2		3	1
u_3	2	2	

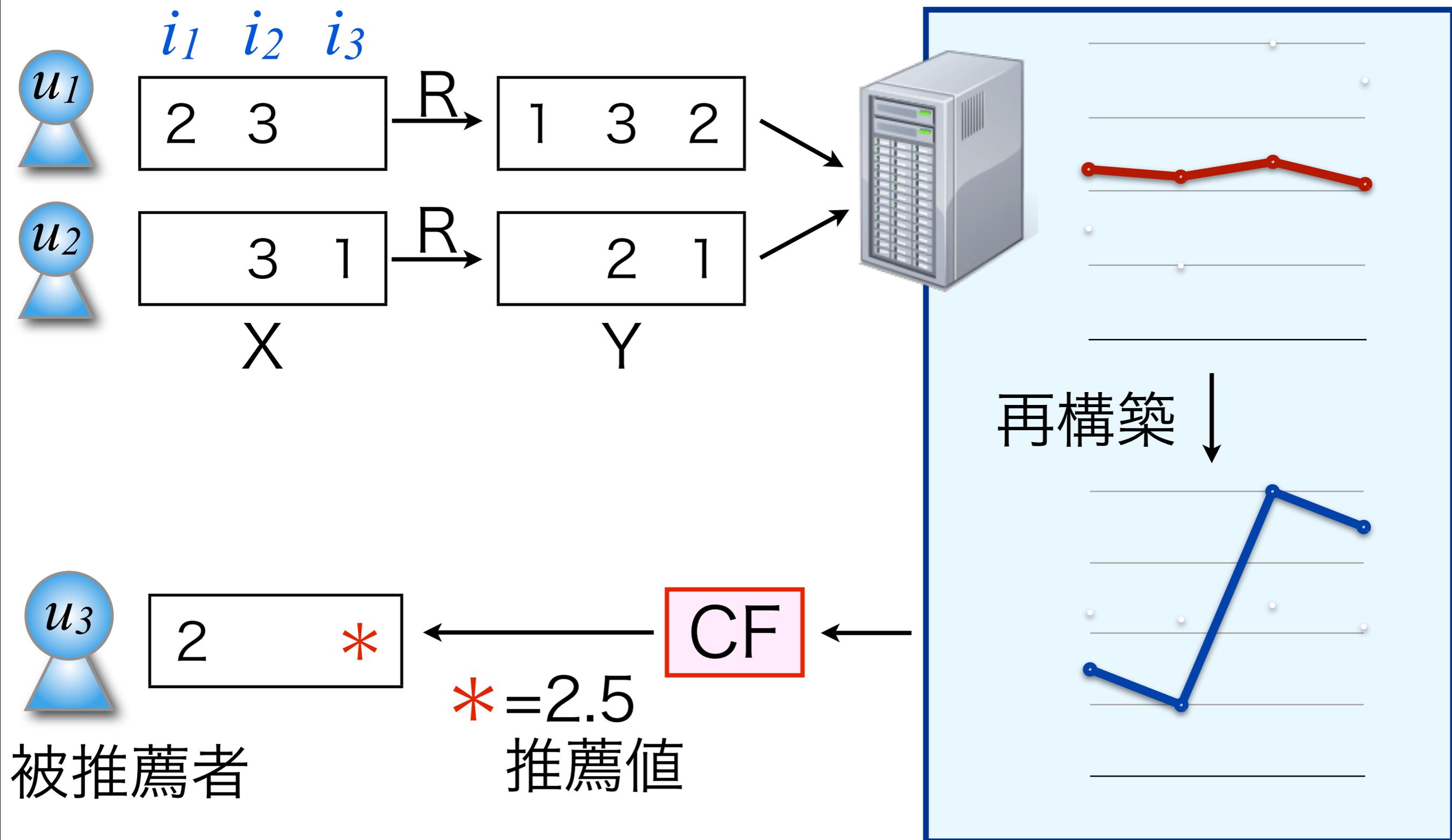
摂動化
→

	i_1	i_2	i_3
u_1	1	3	2
u_2		2	1
u_3	3	1	

X : オリジナル評価値

Y : 摂動化評価値

振動化：再構築 [望月, 菊池 DICOMO2011]



被推薦者

* = 2.5
推薦値

❖ 精度

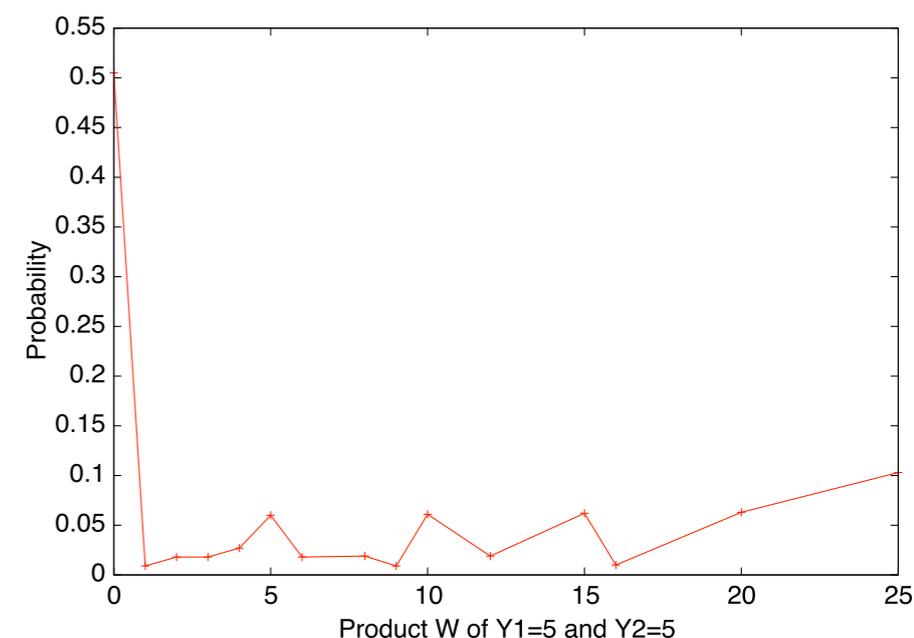
([4] 平均誤差 MAE=1.01)

❖ 原因

(1) CFの複雑な類似度の定義

$$s_{1,2} = \frac{r_{11} \cdot r_{12} + r_{21} \cdot r_{22}}{\sqrt{r_{11}^2 + r_{12}^2} \sqrt{r_{21}^2 + r_{22}^2}}$$

(2) 欠損値の扱い



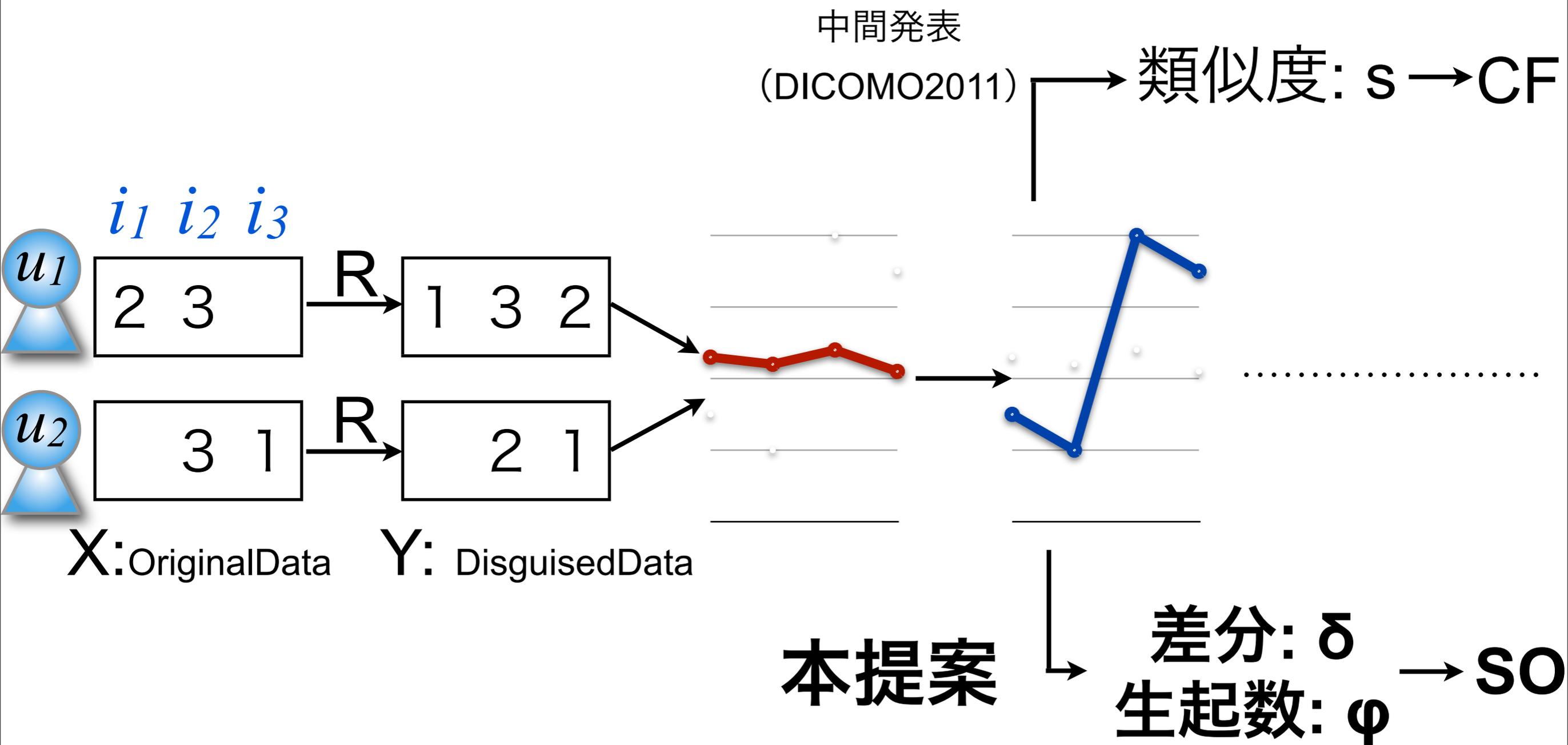
本研究のアプローチ

❖ 摂動化に向けた推薦アルゴリズム **Slope One** 適用

[2]Leniel & Maclachlan, "Slope One Predictors for Online Rating-Based Collaborative Filtering", 2005

	Collaborative Filtering	Slope One
推薦原理	類以度(S)・重み付平均和	平均差分(δ)・共生起数(ϕ)
特徴量	$s_{1,2} = \frac{r_{11} \cdot r_{12} + r_{21} \cdot r_{22}}{\sqrt{r_{11}^2 + r_{12}^2} \sqrt{r_{21}^2 + r_{22}^2}}$	$\delta_{1,2} = (r_{11} - r_{12}) + (r_{21} - r_{22})$
推薦予測	$\frac{\bar{r} + \sum s \cdot r}{\sum s}$	$\frac{\sum (\delta + \phi r)}{\sum \phi}$

提案方式



Slope One 推薦方式

	i_1	i_2	i_3
u_1	2	2	
u_2		3	1
u_3	2		3

u_4	3	2	*
-------	---	---	---

$\phi_{i,3}$ 差分	-1	2	-
--------------------	----	---	---

$\delta_{i,3}$ 生起数	1	1	-
-----------------------	---	---	---

$$* = \frac{(\delta_{13} + r_{41} \cdot \phi_{13}) + (\delta_{23} + r_{42} \cdot \phi_{23})}{\phi_{13} + \phi_{23}}$$

$$= \frac{(1 + 3 \cdot (-1)) + (1 + 2 \cdot 2)}{(-1) + 2} = 3$$

欠損数の予測

未知

オリジナルデータ : X

$N = 16$

	i_1	i_2	i_3	i_4
u_1	■	□	■	□
u_2	■	■	□	■
u_3	□	■	■	□
u_4	□	□	■	□

欠損数 $K_X = 8$

可観測

摂動化データ : Y

	i_1	i_2	i_3	i_4
u_1	■	□	■	□
u_2	■	■	□	□
u_3	□	□	■	□
u_4	□	■	■	■

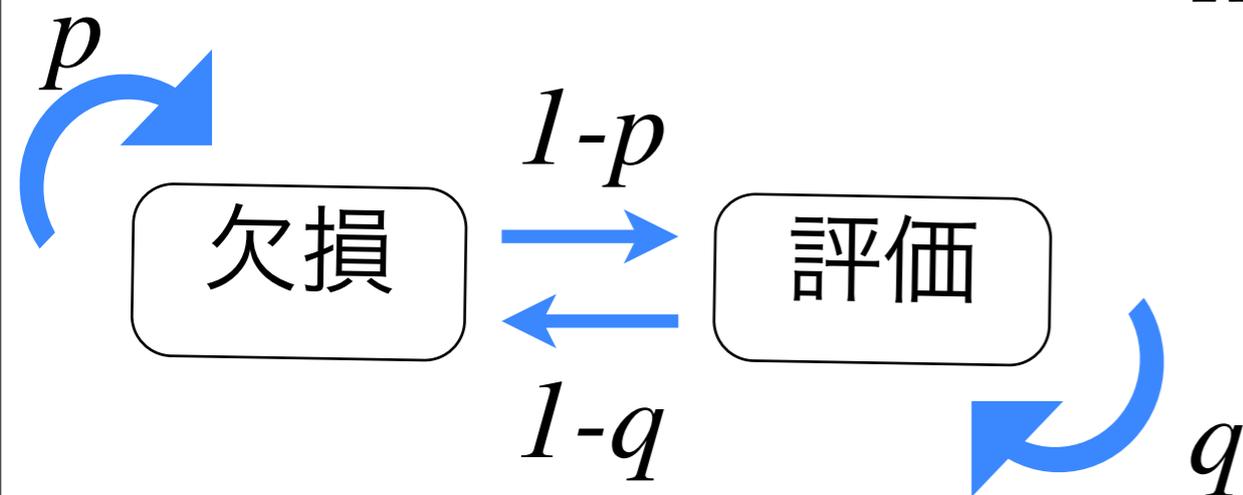
$K_Y = 8 - 2 + 2 = 8$

摂動化

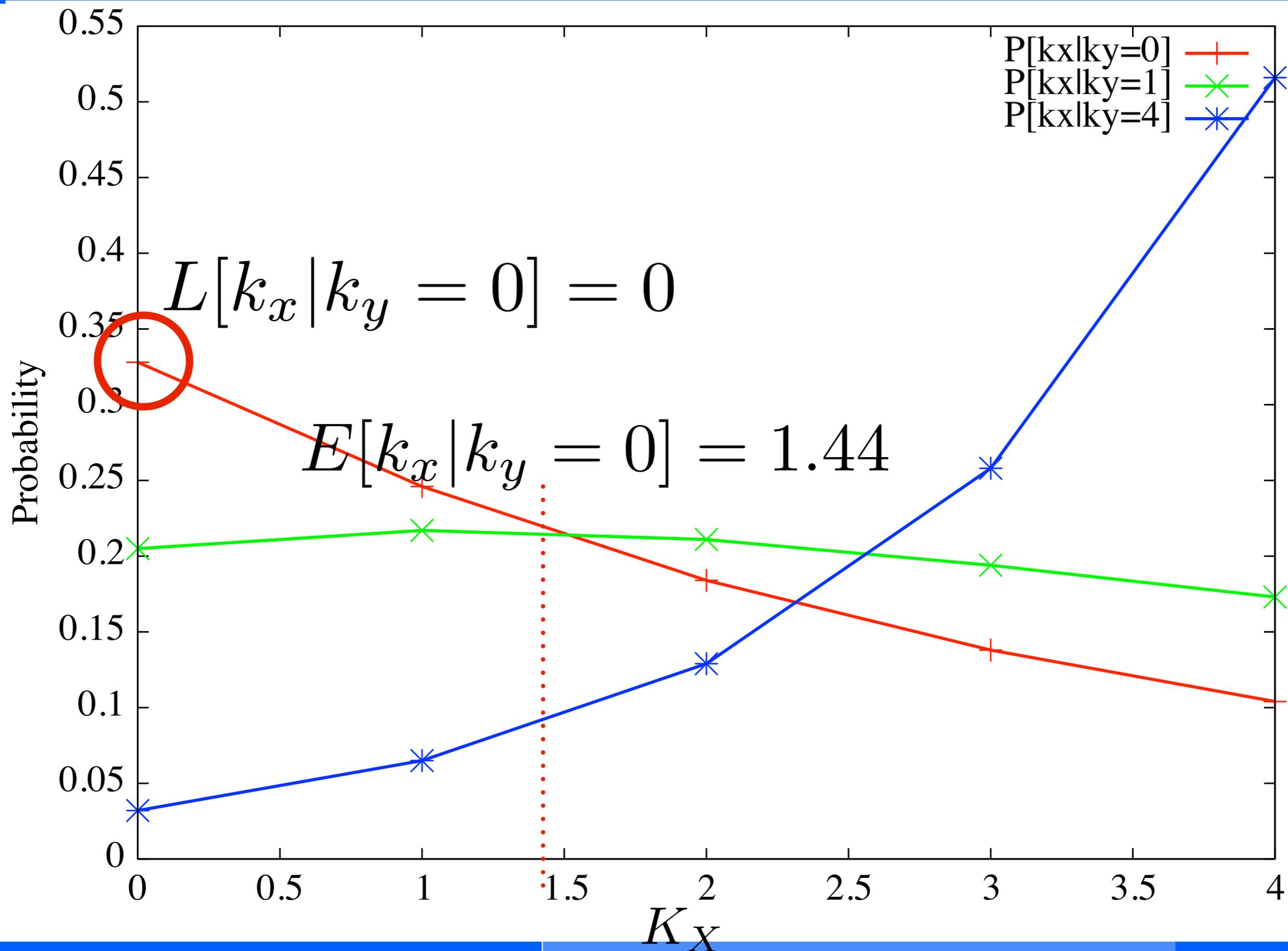
$p = 3/4$

$$K_Y = K_X - K_X(1 - p) + (N - K_X)(1 - p)$$

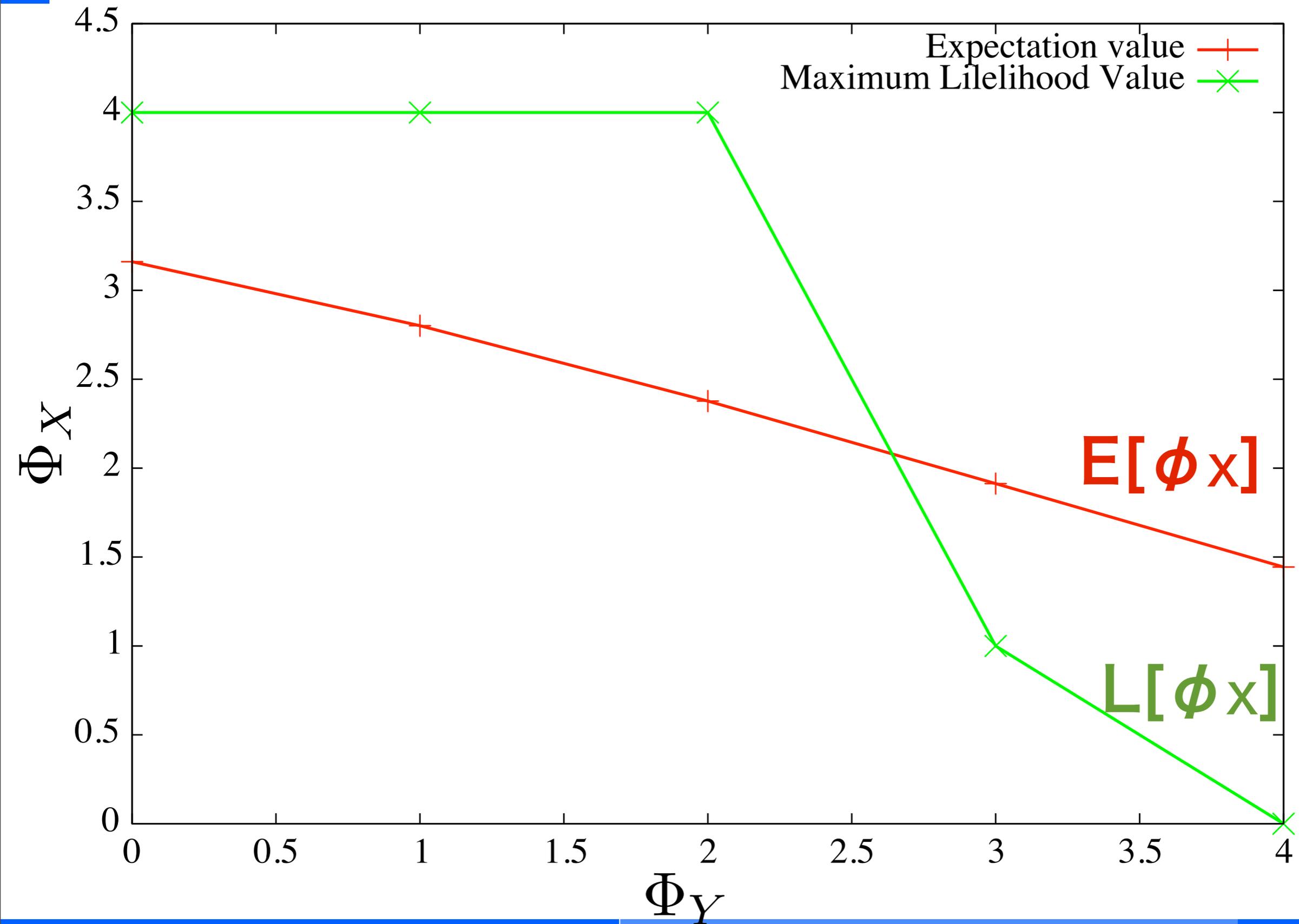
$$= L[K_Y | K_X]$$



$P(K_x|K_y)$ の確率分布



ϕ の推薦値



アイテム間差分 δ の推定

オリジナル：X

i	j	$\delta_{i,j}$
3	2	+1
1	1	0
2	1	+1
2	3	-1

$$\Delta_X = +1$$

摂動化：Y

i	j	$\delta_{i,j}$
3	3	0
2	1	+1
2	1	+1
2	3	-1

$$\Delta_Y = +1$$

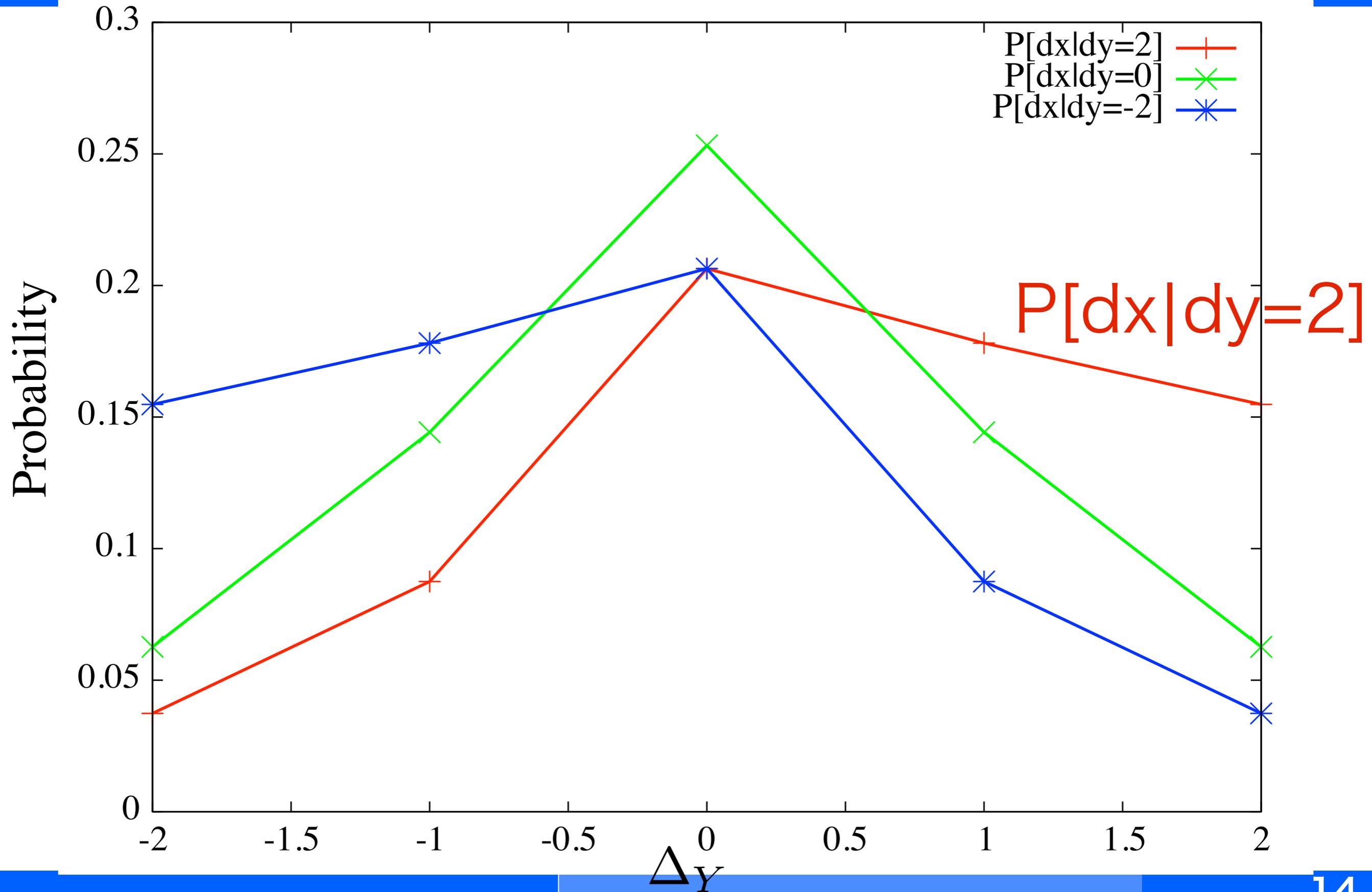
摂動化

← ベイズ推定

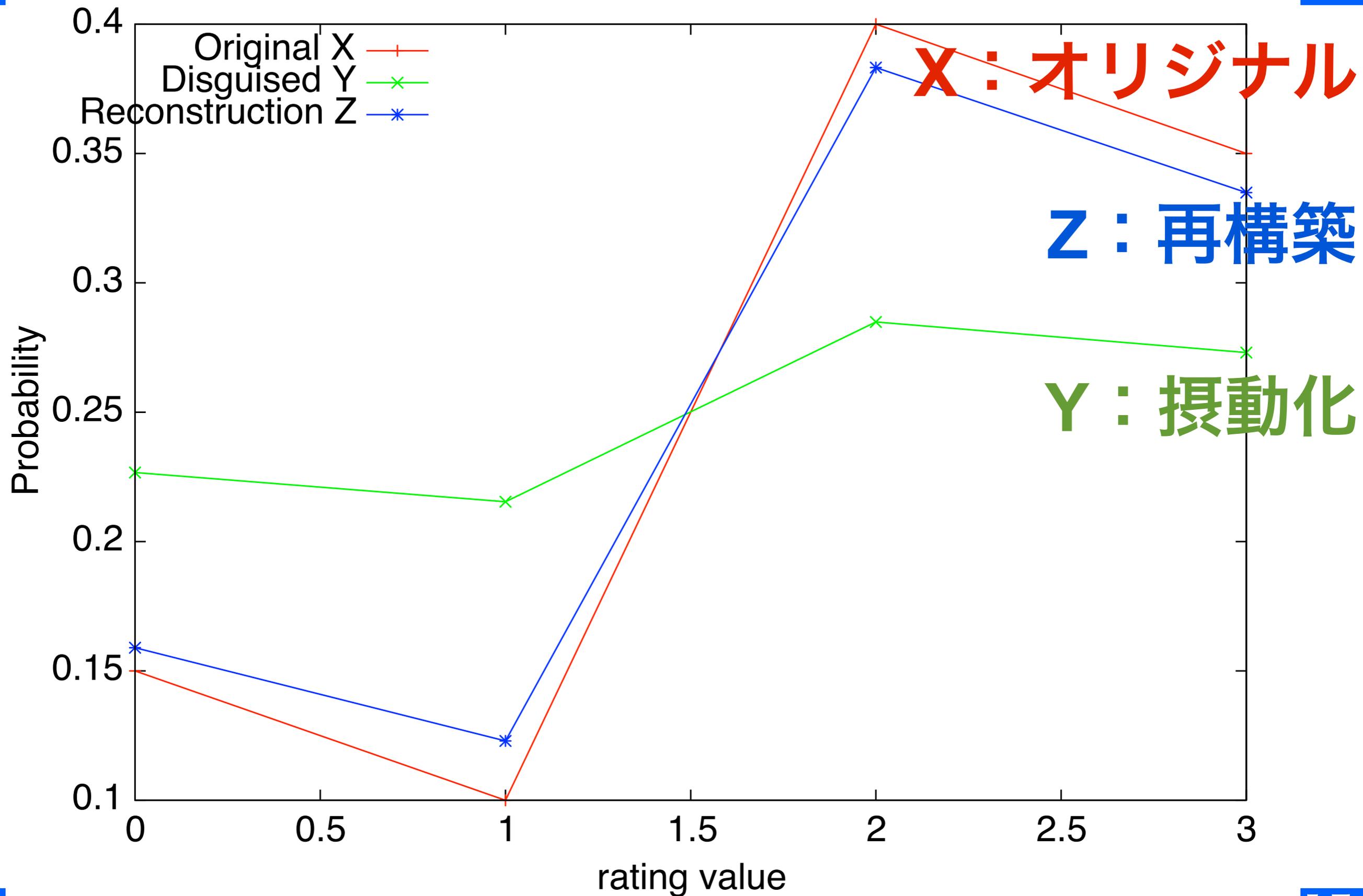
$$\Delta_X = L[\Delta_X | \Delta_Y]$$

$$\begin{aligned} P(\Delta_X | \Delta_Y) &= \sum P(X_1, X_2 | \Delta_Y) \\ &= \sum P(X_1 | \Delta_Y) \cdot P(X_2 | \Delta_Y) \end{aligned}$$

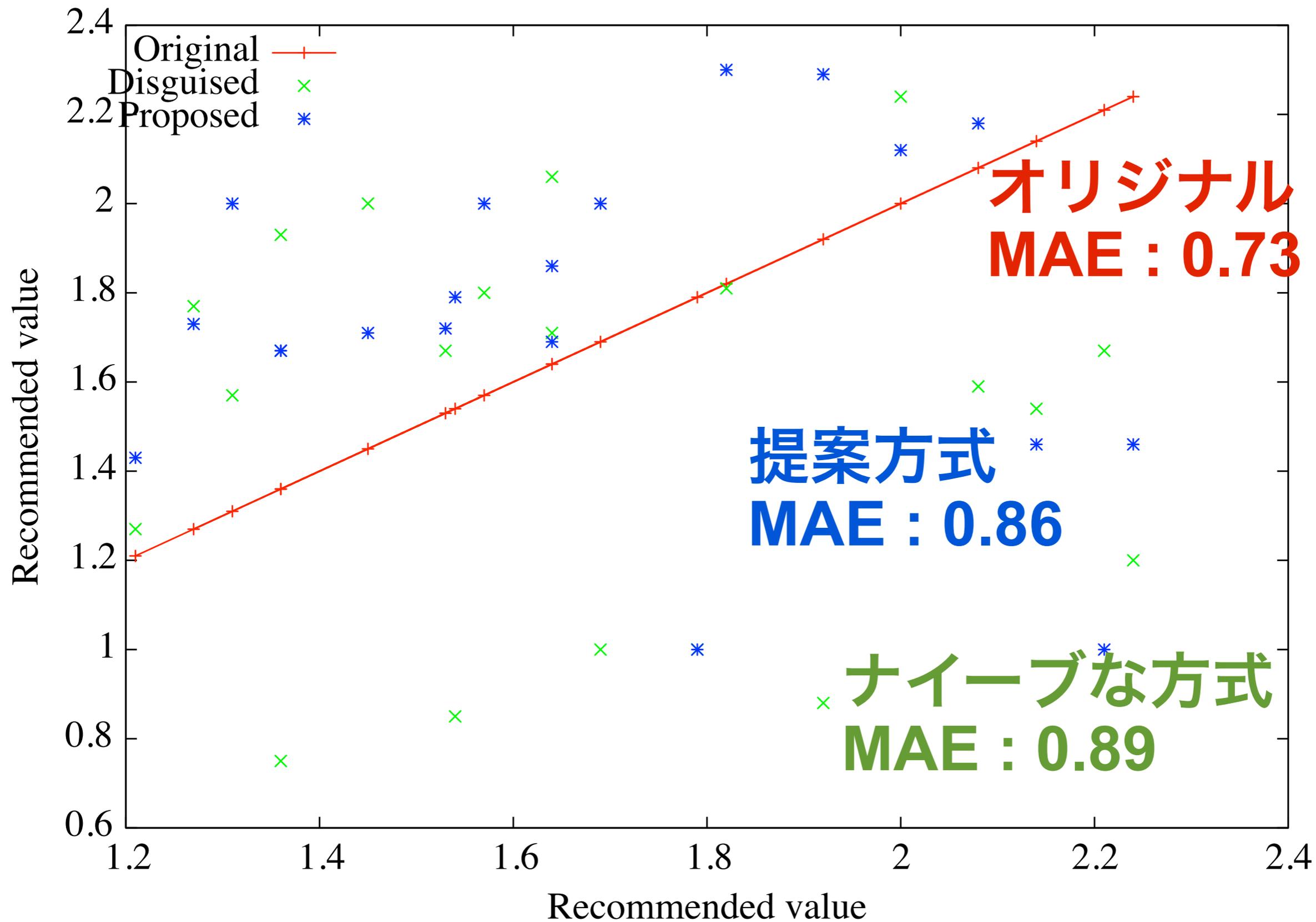
差分 δ の確率分布



再構築精度



Slope One による推薦値の分布



CF とSO の比較

	<i>Collaborative Filtering</i>	<i>Slope One</i>
<i>Original</i>	0.97	0.73
<i>Perturbation</i>	1.03	0.89
<i>Proposed</i>	1.01	0.86

◆まとめ

- 分布のみによる再構築の精度は高い
- Collaborative Filtering に比べ, Slope One は誤差が少ない

◆今後の課題

- 精度の向上
- アイテム依存維持確率の摂動化