

「計量心理学」 落穂拾い

ベイズ・数量化・信頼性

岡本安晴

2018年3月

於：日本女子大学西生田キャンパス

計量心理学とは

田中良久・中谷和夫（1969） 第1章 序論 (p. 1)

田中良久（編）「講座心理学 2 計量心理学」、東京大学出版会

計量心理学（psychometrics）は、心理学研究法の重要な部分をなす測定法、ならびに数量化を研究するものであって、大ざっぱに言えば精神測定法ないしは心理学的測定法をその主要課題とする。しかし、たんに、これら数量的研究法のみを取り扱うものではなく、心理学における数量化の基礎を取り扱うものである。それはたんにある議論またはモデルによる数量化を問題にするだけでなく、実際の測定方法の検討、さらにこれらの方法を心理学的事象に適用した場合に生ずる種々な方法論上の問題をも取り扱う。したがって計量心理学は、一方では心理学研究法と連関し、その主要部分をなすとともに、他方では数理心理学とも連関し、心理学的事象に対する数学の適用の基礎を検討するものである。

数学と心理学

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (ed.). *Handbook of experimental psychology*. (pp. 1-49). New York: John Wiley & Sons, Inc.

The stature of a science is commonly measured by the degree to which it makes use of **mathematics**. (p. 1)

Before the paraphernalia of mathematics can be of use to psychology, we must devise scales of **measurement**. (p. 21)

R. Levy & R. J. Mislevy (2016)
Bayesian Psychometric Modeling.
Boca Raton: CRC Press

Psychometrics concerns
the principles and practices
of reasoning in **assessment.**
(p. xix)

ベイズ

最小 2 乗法

データにモデルを合わせる

複雑な構造のデータに対する
統計的検討に難

最尤法

フィッシャー情報量

漸近定理、
独立に同じ分布に従う多数のデータ

ベイズ統計法

IRTにおいて有用性を確認
e.g., Patz & Junker (1999)

数量化

調査祇

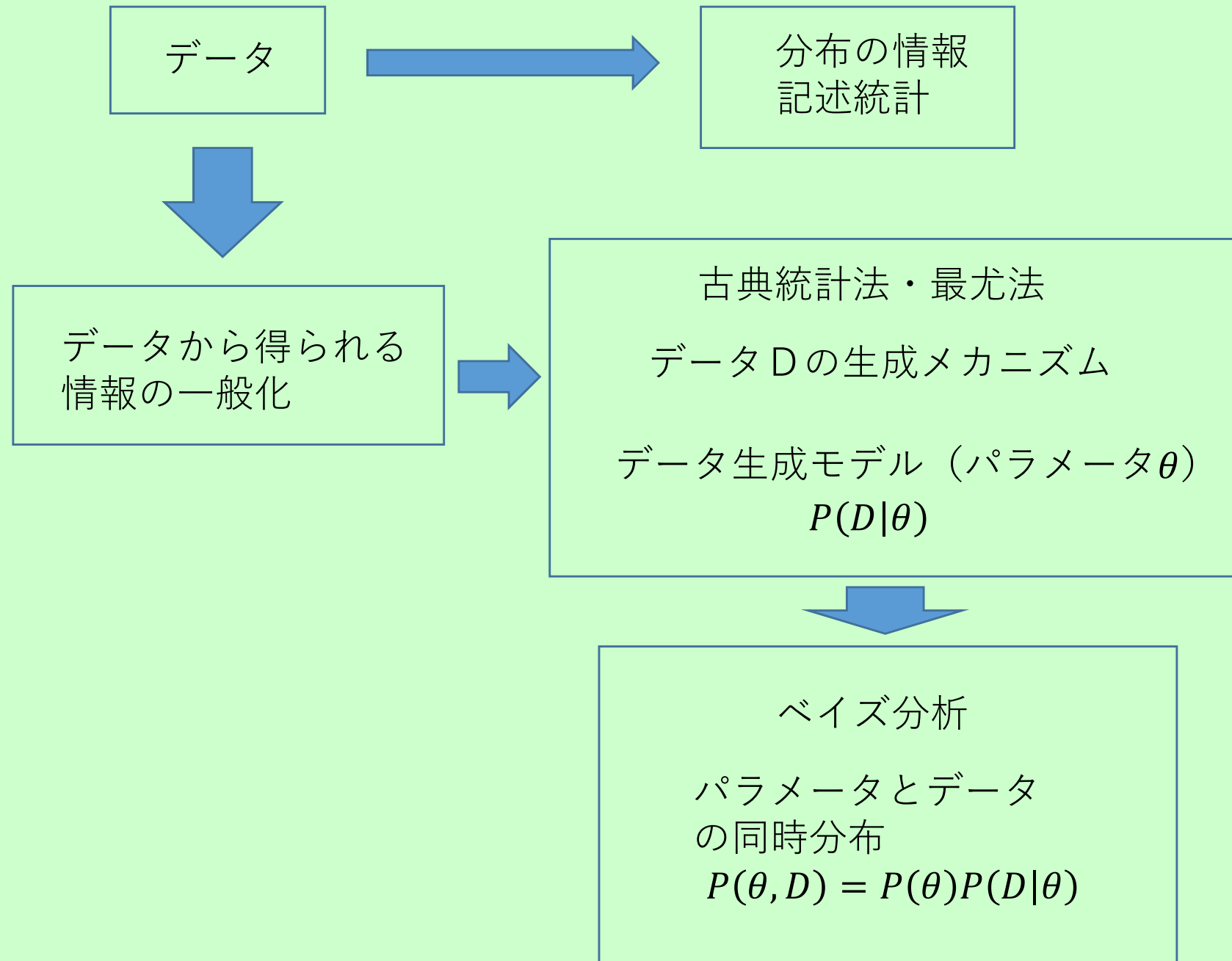
カテゴリ (n 件法、間隔尺度)
カテゴリ (名義尺度)

信頼性

古典テスト理論

連続値項目

カテゴリ値項目に対しては？



分析モデルの進化

豊かな

分析が

容易に

なる

D データ

$P(D|\theta)$

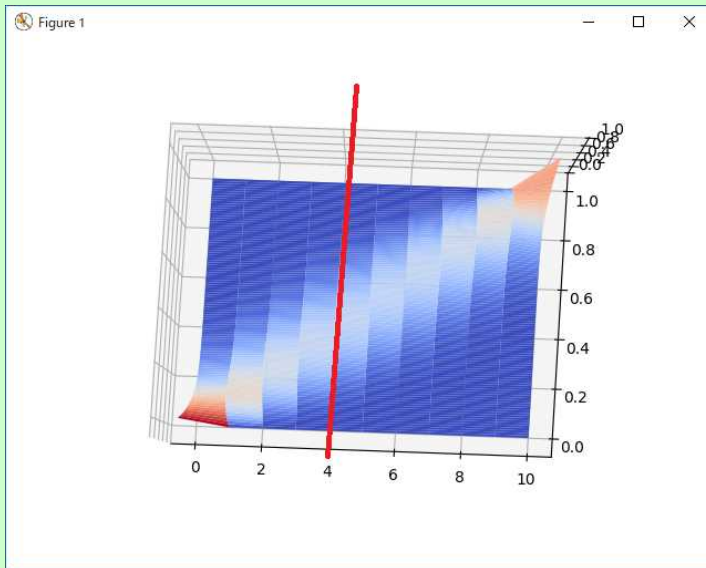
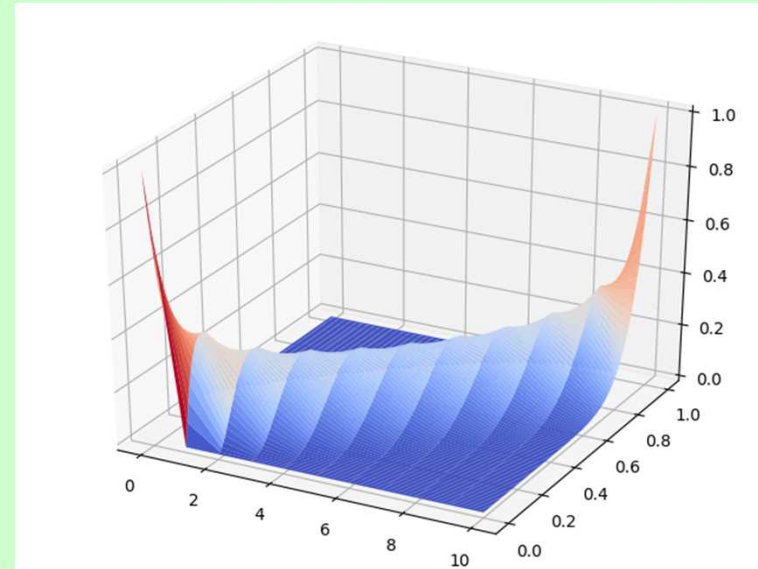
データ生成モデル： θ 固定

$P(D, \theta) = P(D|\theta)P(\theta)$

データとパラメータの同時分布



$$P(p, X) = \binom{10}{X} p^X (1 - p)^{10 - X}$$



$$X=4$$

上記グラフの描画Pythonスクリプト

<http://y-okamoto-psy1949.la.cocacn.jp/Python/misc/BysnBinDistri/>

| | + | - |
|------|-----|-----|
| 該当者 | 0.9 | 0.1 |
| 非該当者 | 0.1 | 0.9 |

最尤法

$$L(\text{該当者}|+) = P(+|\text{該当者}) = 0.9$$

$$L(\text{非該当者}|+) = P(+|\text{非該当者}) = 0.1$$

$$L(\text{該当者}|+) > L(\text{非該当者}|+)$$

| | + | - | 事前確率 |
|------|----------------------------------|----------------------------------|------|
| 該当者 | 0.9 $0.9 \times 0.01 = 0.009$ | 0.1 $0.1 \times 0.01 = 0.001$ | 0.01 |
| 非該当者 | 0.1 $0.1 \times 0.99 = 0.099$ | 0.9 $0.9 \times 0.99 = 0.891$ | 0.99 |
| 周辺確率 | $0.009 + 0.099 = 0.108$ | $0.001 + 0.891 = 0.892$ | |

ベイズ法

$$P(\theta, x) = P(x|\theta)P(\theta)$$

事前確率 $P(\theta) = 0.01, 0.99$

事後確率 $P(\theta|x) = P(\theta, x)/P(x)$

$$P(\text{該当者}|+) = \frac{0.009}{0.108} \approx 0.0833$$

$$P(\text{非該当者}|+) = \frac{0.099}{0.108} \approx 0.9167$$

最尤法

木を見て、森を見ず： $L(\theta|x)$

ベイズ法

木を森の中に見る： $P(\theta|x) = \frac{P(\theta, x)}{P(x)}$

森 = 同時確率 $P(\theta, x)$

共通数量化による分析例

心理学科入学者の分析

参考文献

岡本安晴 (2017) 日本行動計量学会第45回大会抄録集、314 – 317.

共通数量化の目的

1. 数量化をより易しく解り易く

Cf. 数量化、双対尺度法、対応分析、
非線型多変量分析 など

2. 複数の分析において共通の数量化

3. 連続量とカテゴリ変量の同時分析

第1段：

変量ごとの共通数量化

第2段：

変量間の関係の分析

| | 項目j | | | |
|---------|-----|------|-----|------|
| 回答カテゴリ | A | B | C | D |
| 数量化 (例) | 0.7 | -0.1 | 0.1 | -0.7 |

C を選択 \Rightarrow 数量化 0.1

$$0.1 = 0 \times 0.7 + 0 \times (-0.1) + 1 \times 0.1 + 0 \times (-0.7) = (0 \quad 0 \quad 1 \quad 0) \begin{pmatrix} 0.7 \\ -0.1 \\ 0.1 \\ -0.7 \end{pmatrix}$$

$$0.1 = g\omega$$

$$g = (0 \quad 0 \quad 1 \quad 0)$$

$$\omega = \begin{pmatrix} 0.7 \\ -0.1 \\ 0.1 \\ -0.7 \end{pmatrix}$$

ω_{jk} : 変数 j のカテゴリ k に割り当てられる数値
 x_{ij} : 回答者 i の変数 j への回答

$$g_{ijk} = \begin{cases} 1 & x_{ij} = k \text{ のとき} \\ 0 & x_{ij} \neq k \text{ のとき} \end{cases}$$

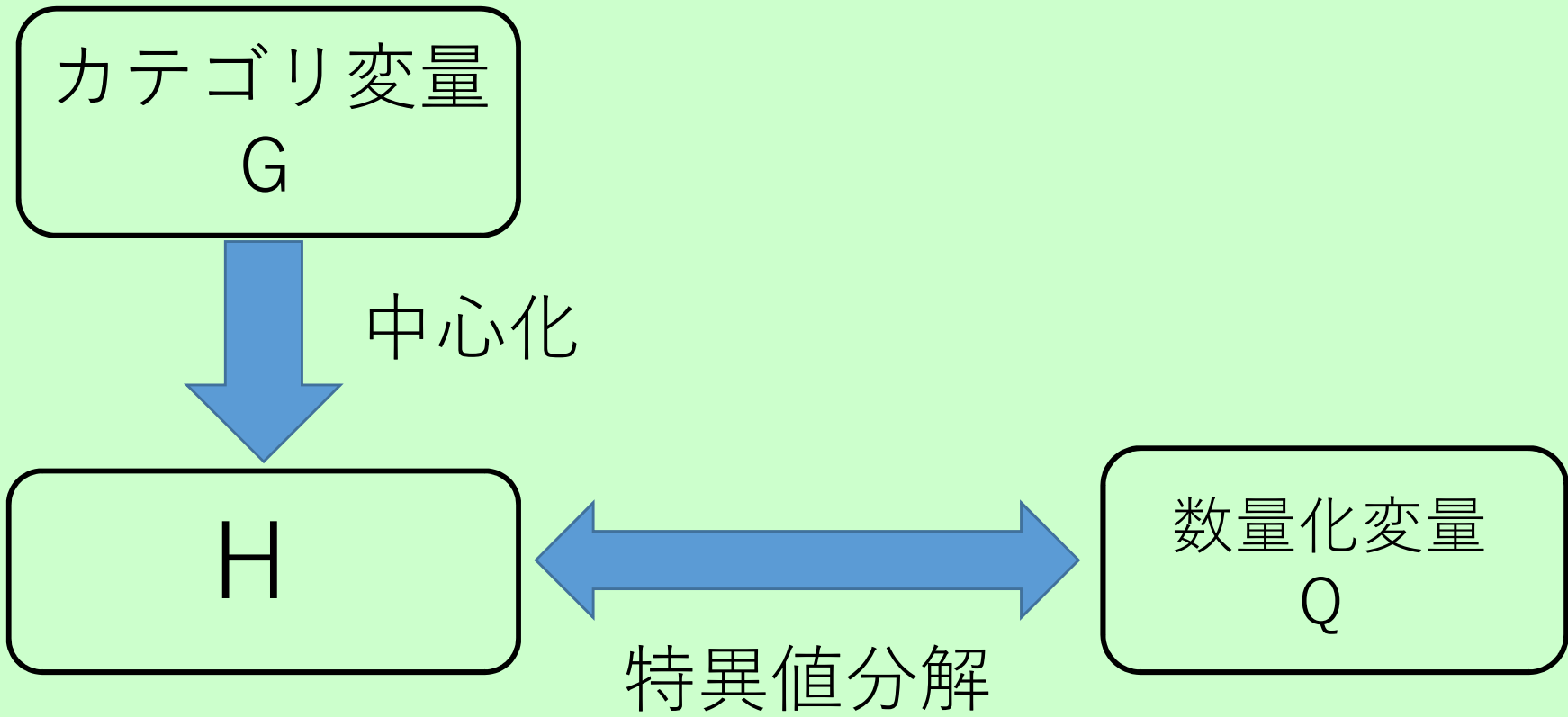
q_{ij} : x_{ij} の数量化された値

$$q_{ij} = \sum_k g_{ijk} \omega_{jk}$$

$$G_j = (g_{ijk}) \quad \omega_j = (\omega_{j1}, \dots, \omega_{jK})'$$

$$Q_j = (q_{1j}, \dots, q_{Nj})' = G_j \omega_j : \text{分散の最大化}$$

$$\omega_j' \omega_j = 1$$



岡本安晴（2017） 共通数量化の応用例. 紀要.

1年生、前期： カテゴリ項目のみ

本報告（2017） 共通数量化の分析例. 学会発表

1年生、後期： カテゴリ項目
+ 連続量（満足度、推薦度）

質問紙調査

以下の問いについて、それぞれ最もよく当てはまるものを選んで（ ）内に○を付けて下さい。

問1 心理学という分野があることを

- () 小学生の頃には知っていた。
- () 中学生の頃に知った。
- () 高校生の頃に知った。
- () その他。

問2 大学受験のとき

- () 心理学≒臨床心理学と思っていた。
- () 心理学には臨床心理学以外の分野があることを知っていた。
- () その他（「心理学とは何をする分野なのか知らなかった。」など）。

問3 大学受験のとき、卒業後は何になりたいと考えていましたか？

- () カウンセラーになりたい。
- () カウンセラーではないが、心理学を活かした職業に就きたい。
- () その他。

問4 いま、卒業後は何になりたいと考えていますか？

- () カウンセラーになりたい。
- () カウンセラーではないが、心理学を活かした職業に就きたい。
- () その他。

問5 自分の得意分野は

- () 文系であると思う。
- () 理系であると思う。
- () その他。

(2枚目へ続く)

問6 心理学科に入学したことにどの程度満足していますか。非常に満足しているなら100点、全然満足していないなら-100点、どちらでもないなら0点、という具合に、-100点から100点までの数値で心理学科に進学したことの満足度を表すと、あなたの満足度は何点になりますか。

() 点

問7 受験生から、どこに進学するのが良いか相談を受けたとしたら、日本女子大学心理学科をどの程度薦めますか。強く薦めるなら100点、強く反対するなら-100点、薦めも反対もしないなら0点、という具合に、-100点から100点までの数値で、日本女子大学心理学科をどの程度薦めたいか評定して下さい。

() 点

表1 数量化の重み $\omega = (\omega_A, \omega_B, \omega_C)'$ の値。
 ω_A 、 ω_B 、 ω_C は、カテゴリA,B,Cの数量化の値を表す。

| | Q1-0 | Q1-1 | Q2-0 | Q2-1 | Q3-0 | Q3-1 | Q4-0 | Q4-1 | Q5-0 | Q5-1 |
|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| A | -0.408 | -0.707 | -0.619 | -0.532 | -0.349 | 0.738 | -0.099 | 0.810 | 0.801 | -0.160 |
| B | 0.816 | 0.000 | 0.771 | -0.270 | 0.814 | -0.067 | 0.751 | -0.320 | -0.539 | -0.614 |
| C | -0.408 | 0.707 | -0.151 | 0.802 | -0.465 | -0.671 | -0.653 | -0.491 | -0.262 | 0.773 |

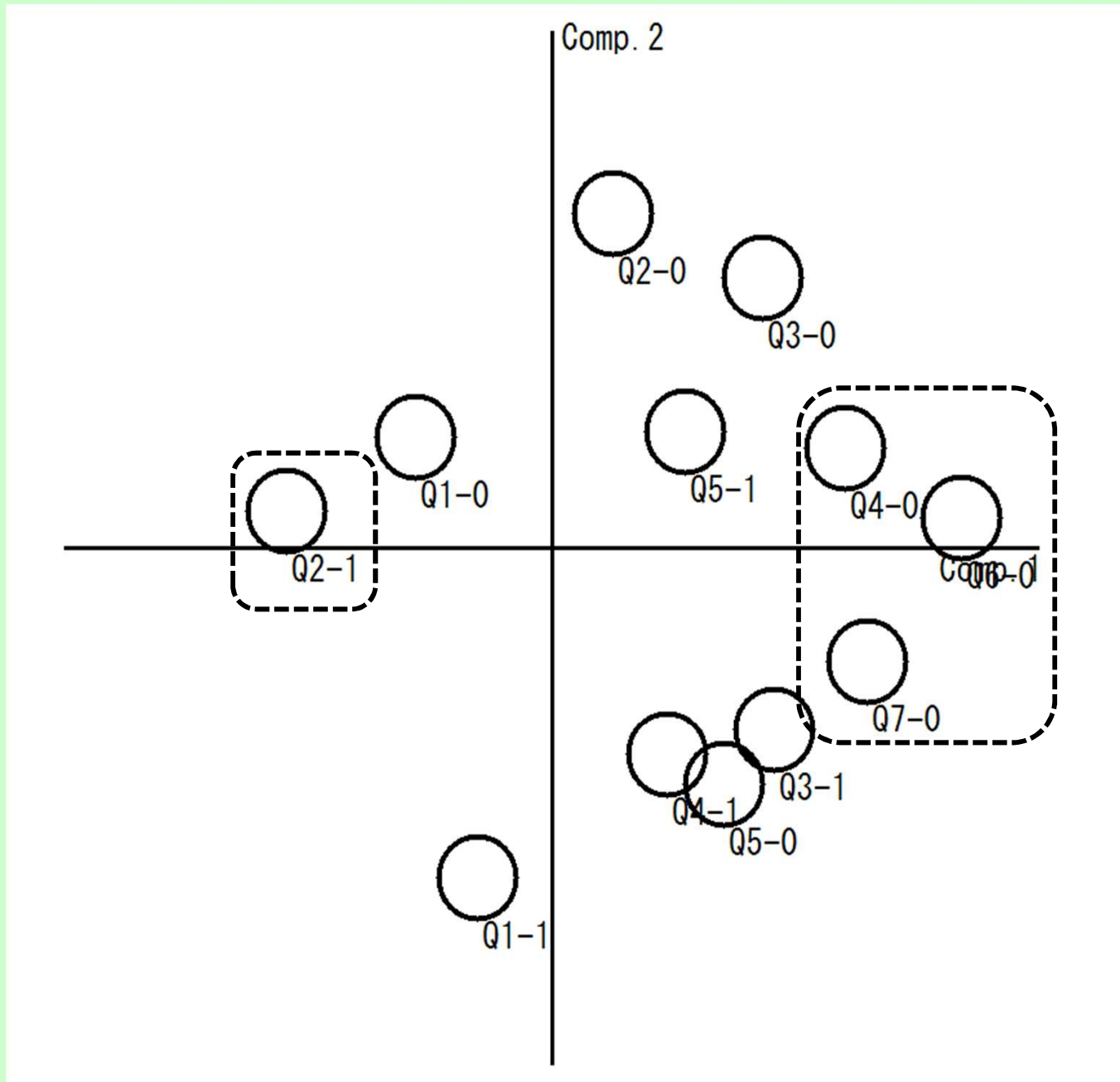


図1 主成分分析による布置

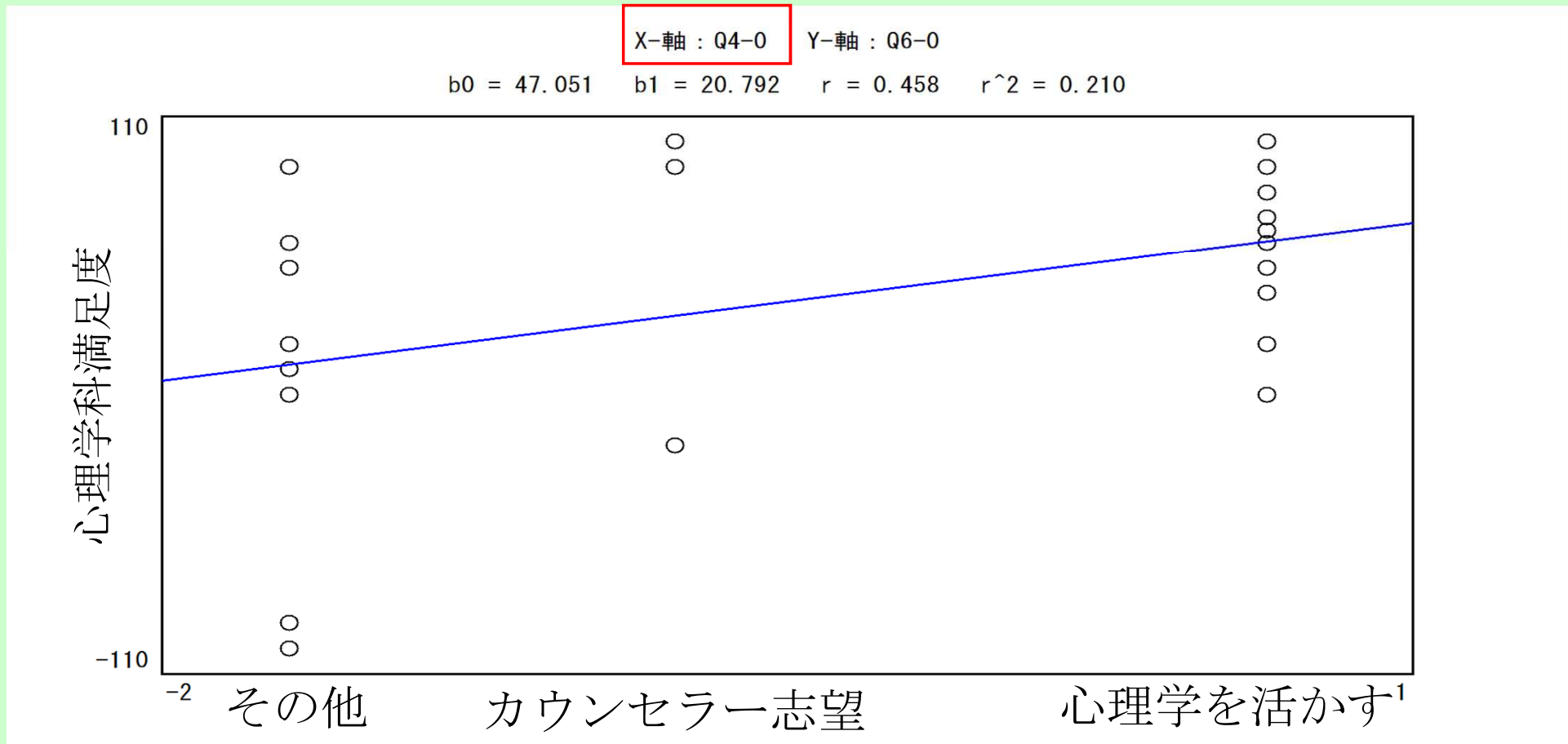


図2 単回帰分析。横軸「いま、卒業後は何になりたいか」と縦軸「心理学科進学への満足度」。

X-軸 : Q2-1

Y-軸 : Q6-0

$b_0 = 47.051$

$b_1 = -19.395$

$r = -0.427$

$r^2 = 0.183$

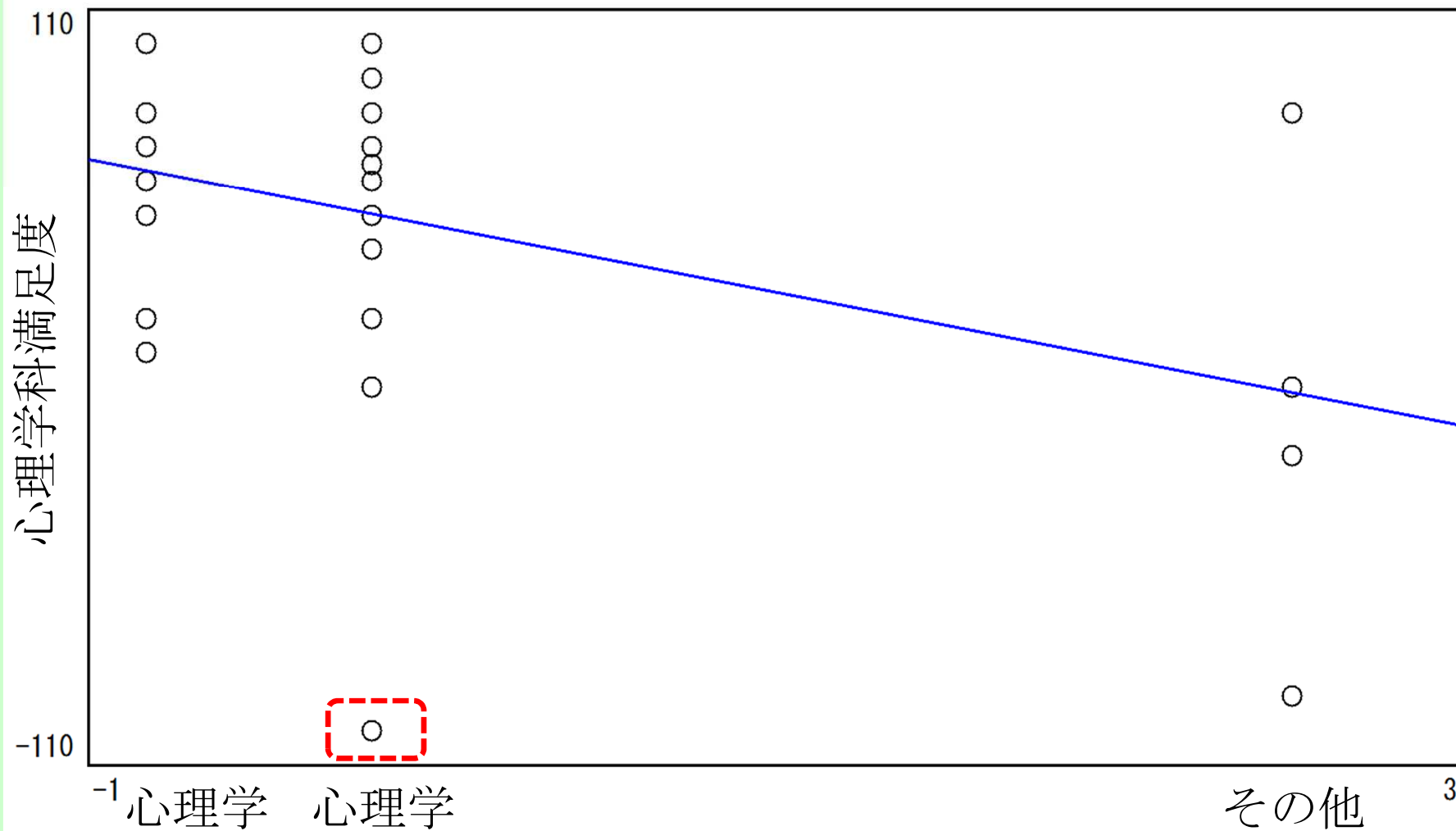


図3 単回帰分析。横軸「大学受験時の心理学のイメージ」と縦軸「心理学科進学への満足度」。

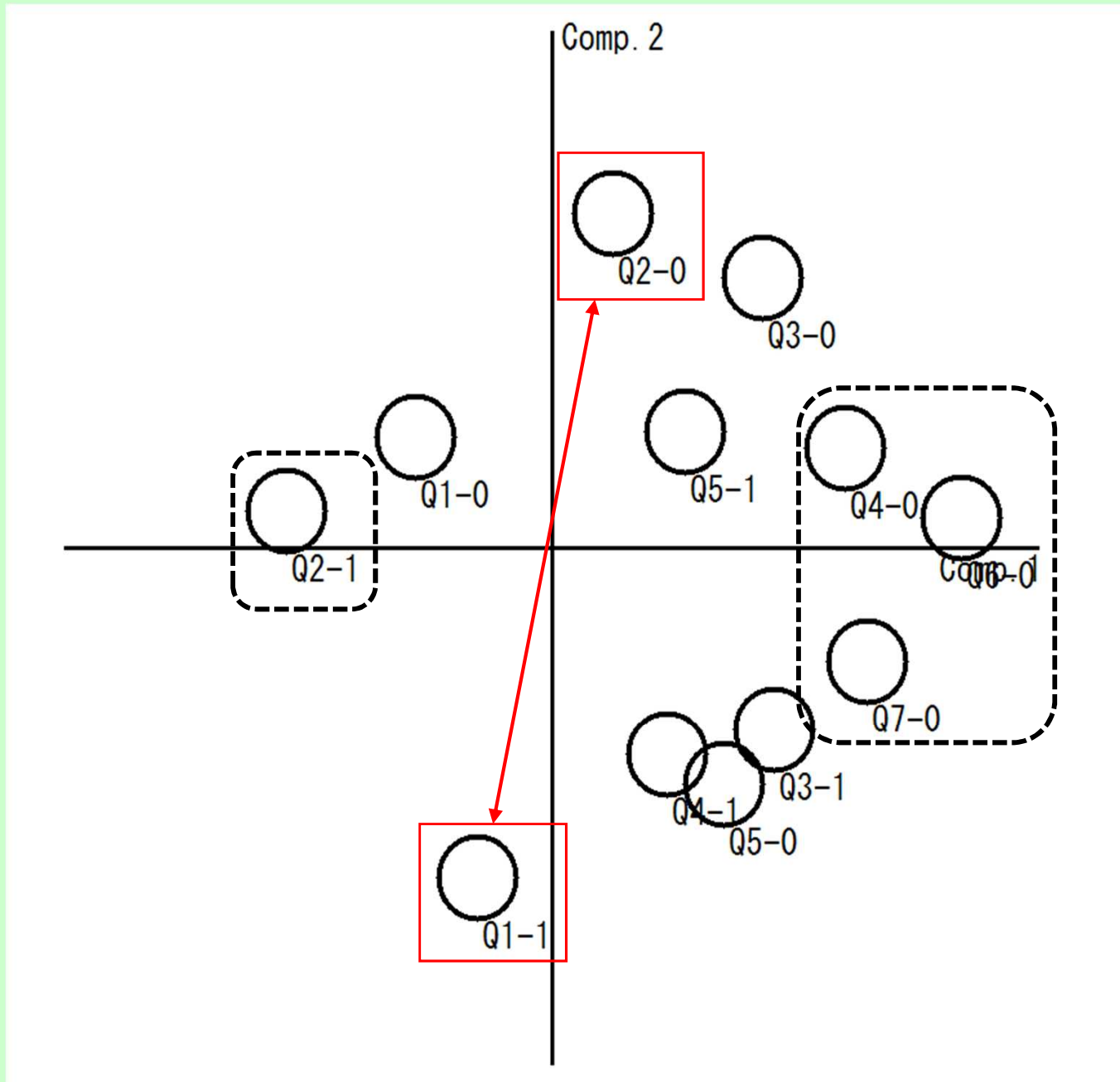


図1 主成分分析による布置

問1と問2のクロス表

| | | 問2 大学受験のとき | | | |
|-------------------------------|---------------|-------------------------|---------------|------------------------|----|
| | | 心理学≒ 臨床心理学 -0.619 | その他 -0.151 | 心理学> 臨床心理学 0.771 | |
| 問1 心理学という分野があることを 知ったとき | 高校生 0.707 | 6 | 2 | 4 | 12 |
| | 中学生 0.000 | 3 | 2 | 10 | 15 |
| | 小学生 -0.707 | 1 | 1 | 10 | 12 |
| | | 10 | 5 | 24 | |

まとめ

1. カテゴリ変量ごとの数量化は簡単である。
2. カテゴリ変量を数量化することにより、量的変量と一緒に量的分析（主成分分析、回帰分析など）が可能になる。
3. カテゴリ変量の数量化により、カテゴリ変量と量的変量の関係が解り易くなる。
4. 関係のあるカテゴリ変量の組を容易に見つけることができるようになる。

項目反応理論における 信頼性の3つの観点

Precision, Consistency, Predictive Power

参考文献

Y. Okamoto (2017).

Three Aspects of Reliability in Item

日本テスト学会第15回大会発表論文抄録集、152 – 155.

測定の基礎モデル

$$\text{測定値} = \text{真値} + \text{誤差}$$

信頼性

Precision
Power

Consistency

Predictive

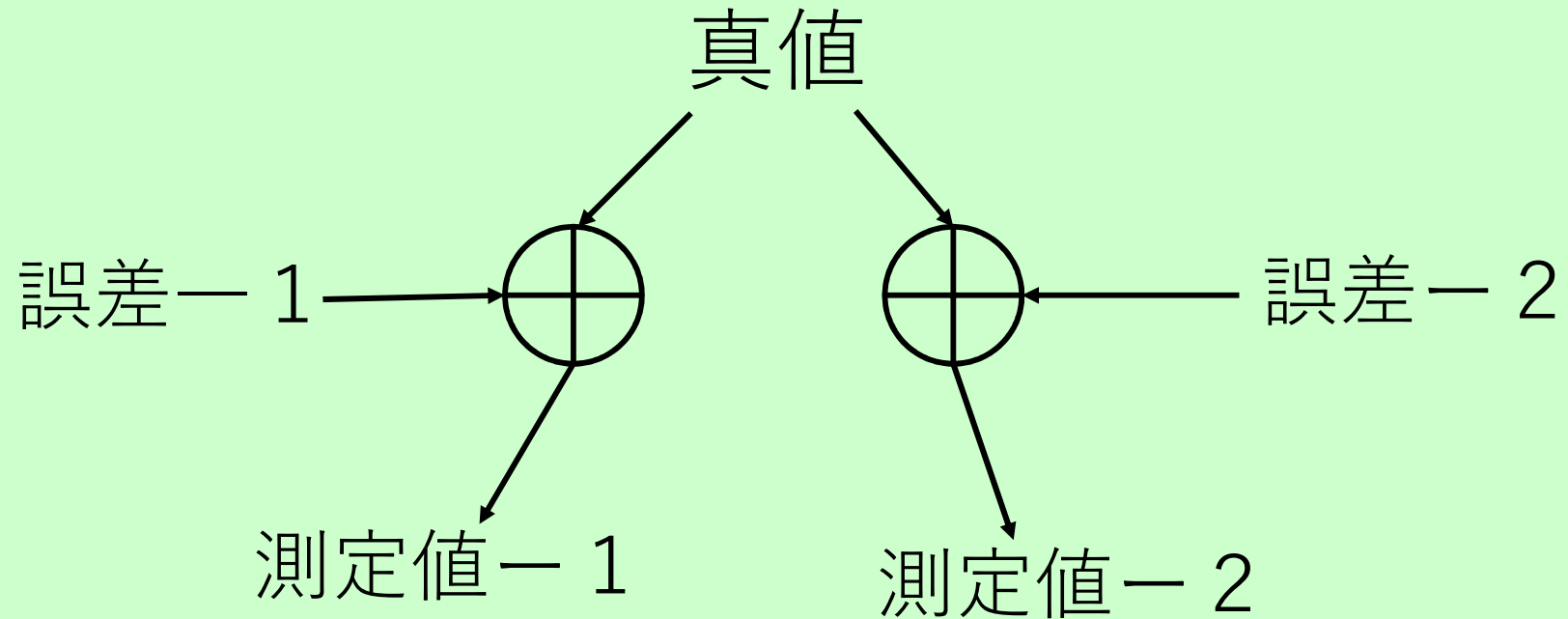
Precision

$$\frac{\text{Var}(\text{真値})}{\text{Var}(\text{測定値})} = \frac{\text{Var}(\text{測定値} - \text{誤差})}{\text{Var}(\text{測定値})}$$

Crocker & Algina (1986)

McDonald (1999)

Consistency



$Cor(\text{測定値} - 1, \text{測定値} - 2)$

Cronbach (1961)

Predictive Power

測定値と真値の単回帰モデル



決定係数

Raykov & Marcoulides (2011)

$$\{Cor(\text{測定値}, \text{真値})\}^2$$

Lord & Novick (1968)

連続項目の場合

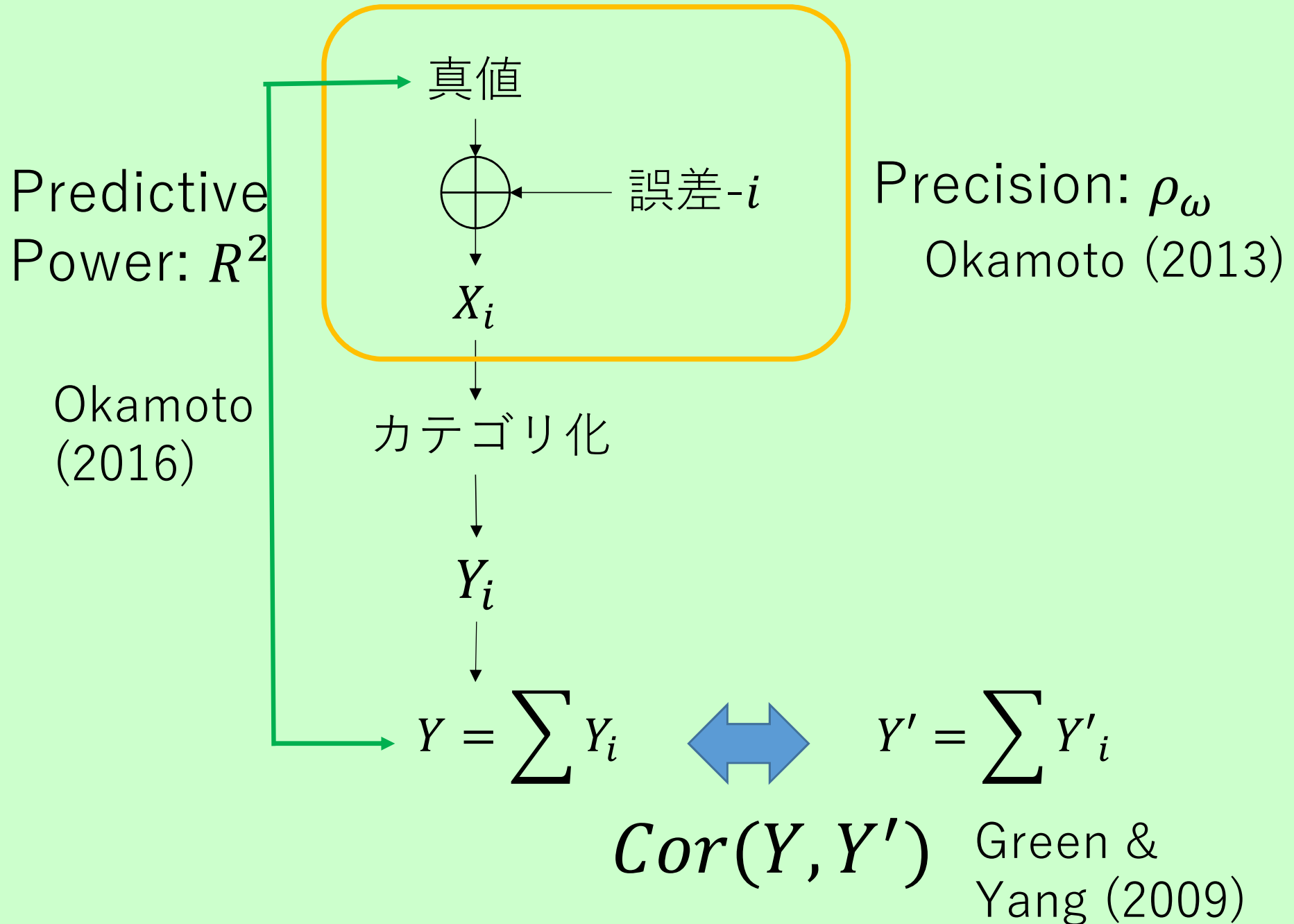
Precision

= Consistency

= Predictive Power

しかし、カテゴリ項目では成り立たない！

カテゴリ項目



カテゴリ項目

Precision: ρ_{ω}

∨

Consistency: $Cor(Y, Y')$

∨

Predictive Power: R^2

項目反応理論の場合は？

項目反応理論（IRT）における 信頼性係数



項目反応理論における
真値と測定値（推定値）



真値とは？

測定対象・概念の値 = 真値

心理学的
(Substantial Science)

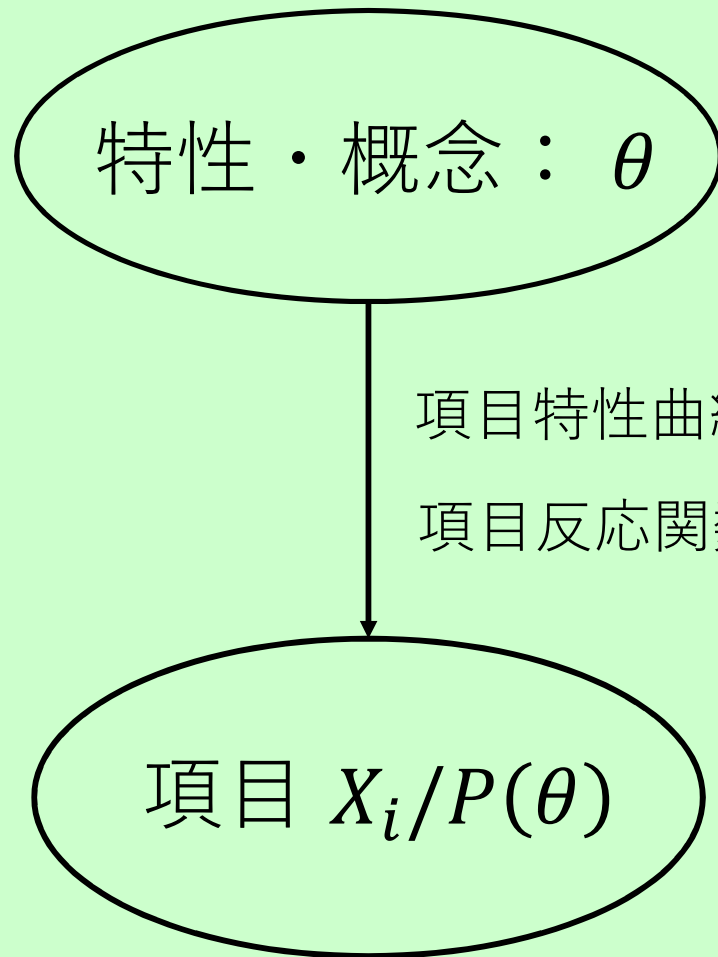
Samejima (1994)
Toyoda (1989)

変数・測定値の平均 = 真値

統計学的
(Formal Science)

Dimitrov (2003)
Kim & Feldt (2010)
Milanzi, Molenberghs,
Alonson, Verbeke &
De Boeck (2014)

項目反應理論



θ ： 測定対象
心理学的真値

$E(X_i)$ ： 平均
統計学的真値

推定値

ベイズ統計法

点推定値／事後分布

事前分布による制約

J. Gill (2015) “Bayesian Methods: A social and behavioral sciences approach, 3rd ed.” p. xxv.

“Bayesian modeling has become standard and MCMC is well-understood and trusted.”

最尤法

フィッシャー情報量／漸近定理

小数データ ⇒ 全問正答／誤答

真値 θ から推定値 $\hat{\theta}$ への図式

X_i : テスト項目 i

一般的には

信頼性

θ_p : 回答者 p の真値

θ_p : カテゴリ可
確率分布可
(連続量/カテゴリ)

$$S(X_{p(i)}) = \{X_{p(1)}, \dots, X_{p(m)}\}$$

$X_{p(j)}$: 回答者 p の第 j 回答項目

$\hat{\theta}_p$: θ_p の推定値

$\hat{\theta}_p$: 点推定値
カテゴリ
確率分布/
(連続量・カテゴリ)

信頼性の3つの観点

Precision

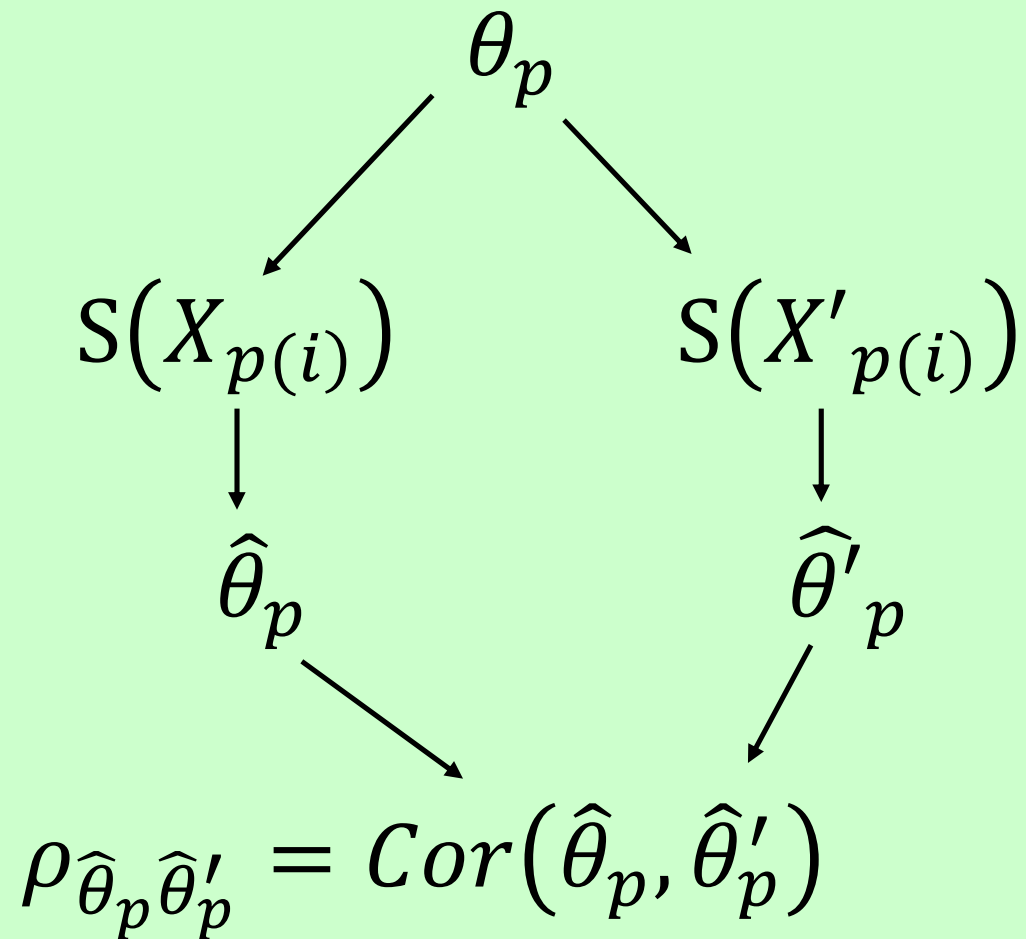
$$\rho_V(\theta_p) = 1 - \frac{V_{post}(\theta_p)}{V_{prior}(\theta_p)}$$

$$\text{cf. } \frac{\text{Var}(\text{真値})}{\text{Var}(\text{測定値})} = 1 - \frac{\text{Var}(\text{誤差})}{\text{Var}(\text{測定値})}$$

$V_{prior}(\theta_p)$: θ_p の事前分布 $N(0, 1)$ の分散

$V_{post}(\theta_p)$: θ_p の事後分布の分散

Consistency/Stability



Predictive Power

単回帰モデル

$$\theta_p = a_1 \hat{\theta}_p + b_1$$

$$\hat{\theta}_p = a_2 \theta_p + b_2$$

共通の決定係数

$$R_{\theta_p \hat{\theta}_p}^2 = \{Cor(\theta_p, \hat{\theta}_p)\}^2$$

シミュレーションによる例

2パラメータ ロジスティック モデル

$$P(X_i = 1|\theta) = \frac{1}{1 + \exp(-1.7a_i(\theta - b_i))}$$

$$a_i = 1 \quad b_i = 0$$

$$S(X_{p(i)}) = \{X_1, \dots, X_M\}$$

Prior distribution for $\theta_p \sim N(0, 1)$

Precision

$$\theta_p = -3, -2, -2, 0, 1, 2, 3 \quad 20 \text{ experiments}/\theta_p$$

$$M = 5$$

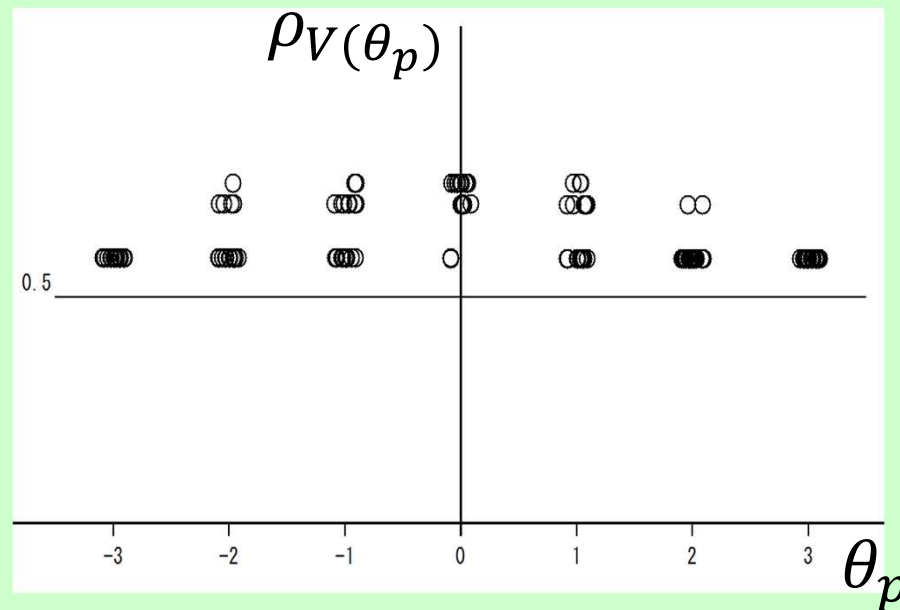


Figure 1. Scattergram of points $(\theta_p, \rho_V(\theta_p))$ for $M = 5$

Precision

$$\theta_p = -3, -2, -2, 0, 1, 2, 3 \quad 20 \text{ experiments}/\theta_p$$
$$M = 100$$

- 信頼性に関わる要因
- 項目数
- 反応カテゴリ数
- 項目の多様性
(難易度)

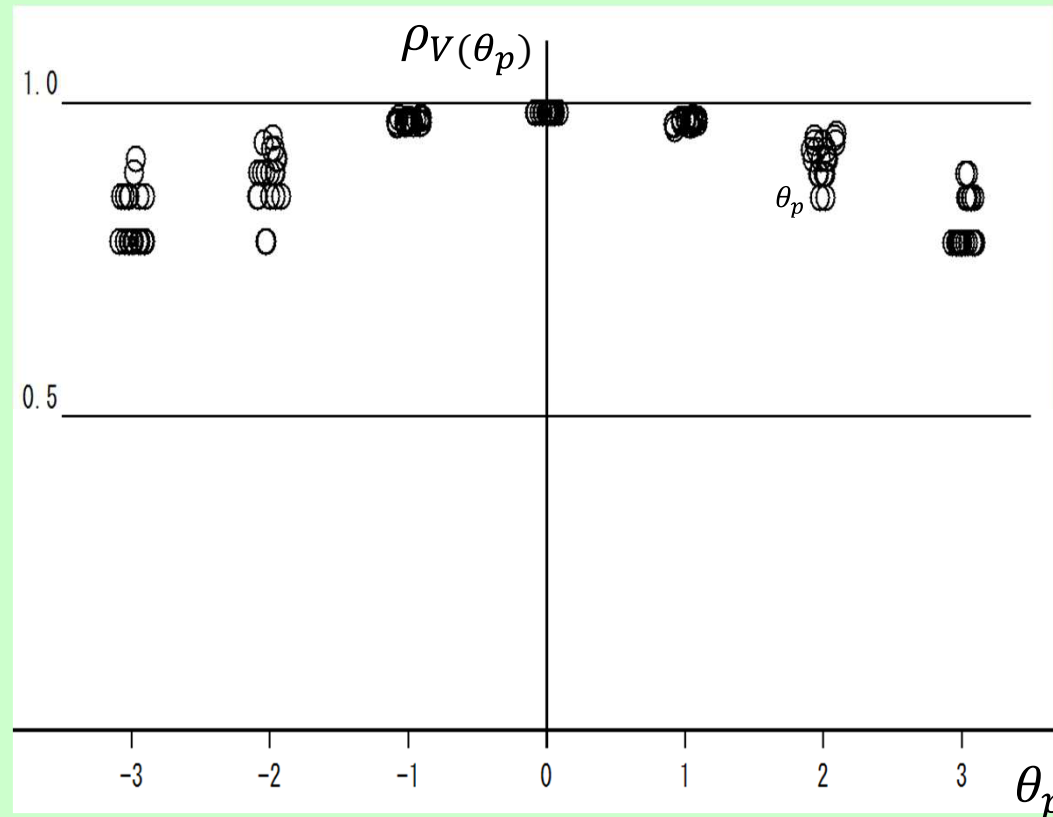


Figure 2. Scattergram of points $(\theta_p, \rho_V(\theta_p))$ for $M = 100$.

Stability/Consistency

1000 samples: $\theta_p \sim N(0, 1)$

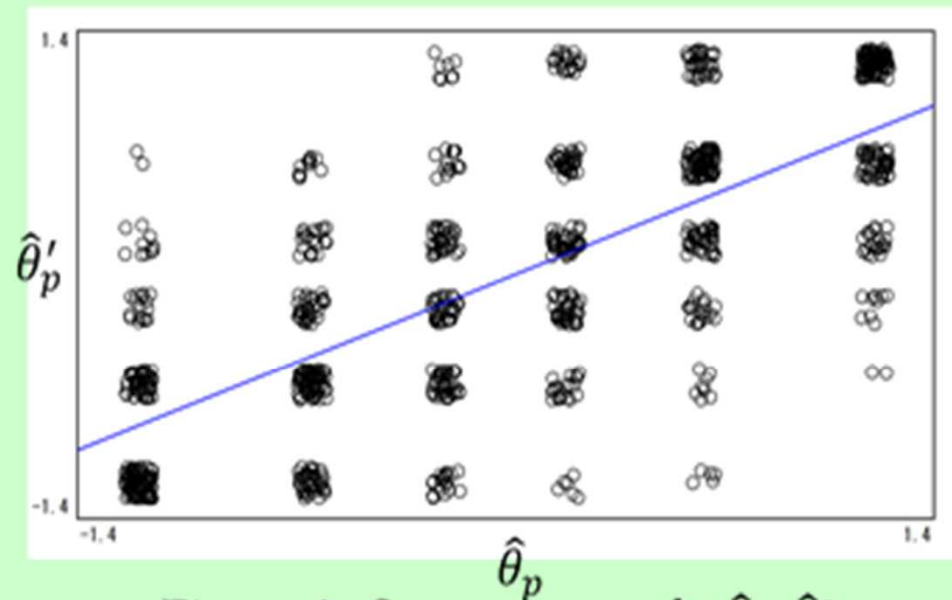
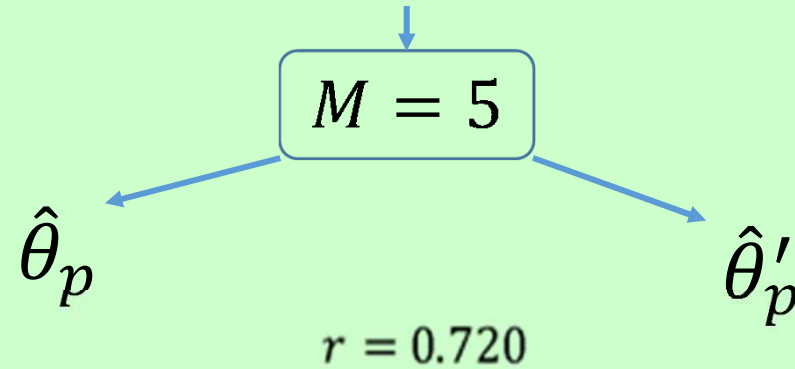


Figure 3. Scattergram of $(\hat{\theta}_p, \hat{\theta}'_p)$

Predictive Power

1000 samples: $\theta_p \sim N(0, 1)$

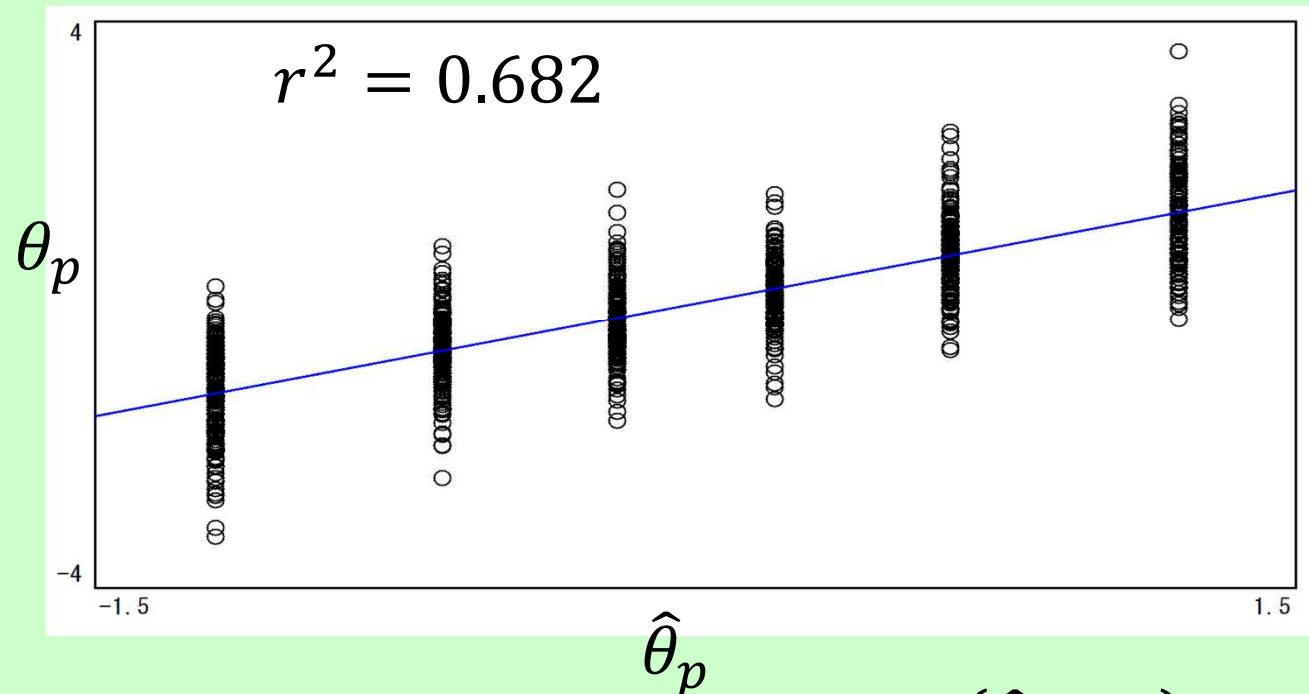
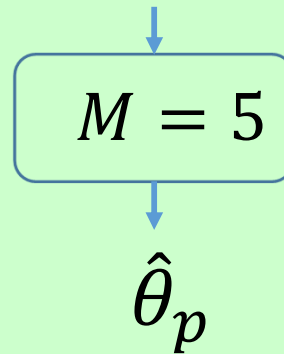


Figure 4. Scattergram of $(\hat{\theta}_p, \theta_p)$.

まとめ

1. ビジュアルに信頼性を表示することにより、推定値の信頼性が直感的にわかる。
2. 真値 θ と推定値 $\hat{\theta}$ との関係として信頼性を考えたとき、真値 θ および推定値 $\hat{\theta}$ は、連続体上の値に限らず、カテゴリ、ランキング上の確率分布として考えることもできる。
3. 上の2の一般化されたモデルにおいても、真値 θ と推定値 $\hat{\theta}$ との関係をシミュレーションによって与える本報告の方法は有用であると考えられる。

まとめ

1. ベイズ統計法は使い易い。
2. カテゴリ変量は、連続変量と統合的に分析可能。
3. 連続変量の理論をカテゴリ変量に適用するのは避けるべき。

それでは