

ブートストラップ法を用いた因子負荷量の推定

- 大学生の結婚に対する意識調査の事例 -

片所 強

目白大学心理学研究科

研究内容の概要

従来的に探索的因子分析を行う `factanal()` を利用

これにより分類された項目の合計点を尺度得点とする

ブートストラップ法を用いる方法

- `factanal()` を用いた探索的因子分析では . . . ×
- `cfa()` を用いた検証的因子分析だと . . .

分類された項目の合計点を尺度得点とする

上記2つの方法によって得られた尺度得点を用いた解析結果を
比較してみる

発表の流れ

調査事例の紹介（使用するデータセット）

心理学における因子分析の役割と使用の問題

ブートストラップとシミュレーション手順

従来法の結果およびBS法による結果の紹介

異なる尺度得点を用いたモデル解析の比較

まとめ・問題点・考察

調査事例の紹介

2006年に大学生を対象に調査を実施

結婚に関する質問

- カテゴリ1 結婚の良い点(10項目)
- カテゴリ2 結婚の悪い点(8項目)
- カテゴリ3 子供を持つことの価値(7項目)
- カテゴリ4 子供を持つことに対する不安(8項目)

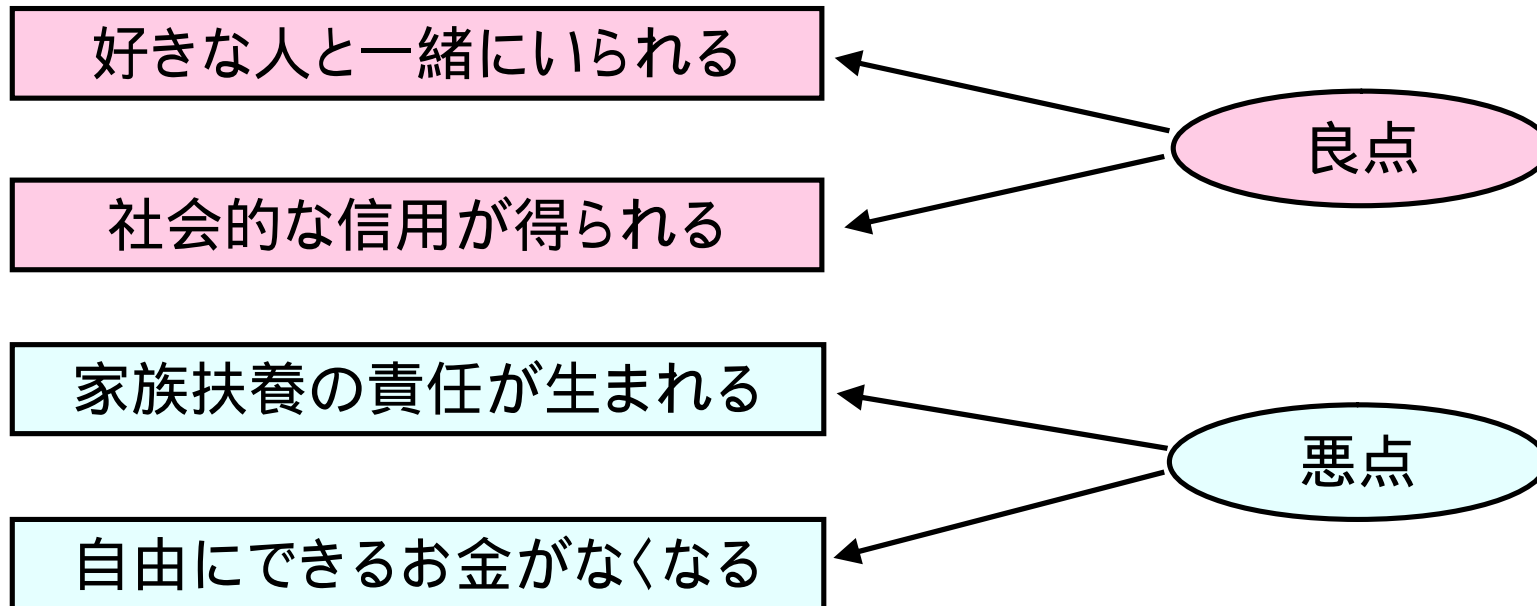
全て30項目

154名の有効回答を得ることができた

心理学における因子分析の役割

質問項目の分類と尺度得点の計算

(例) 結婚の良点に関する項目群と悪点に関する項目群を分類



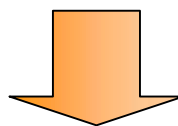
因子分析の使用に関する問題

推定された因子負荷量はどれだけ信頼できるか

回帰分析における回帰係数についての検定・信頼区間

標本(データセット)の変動によって因子構造が変わる

その質問項目は本当にその因子に含まれるべきか

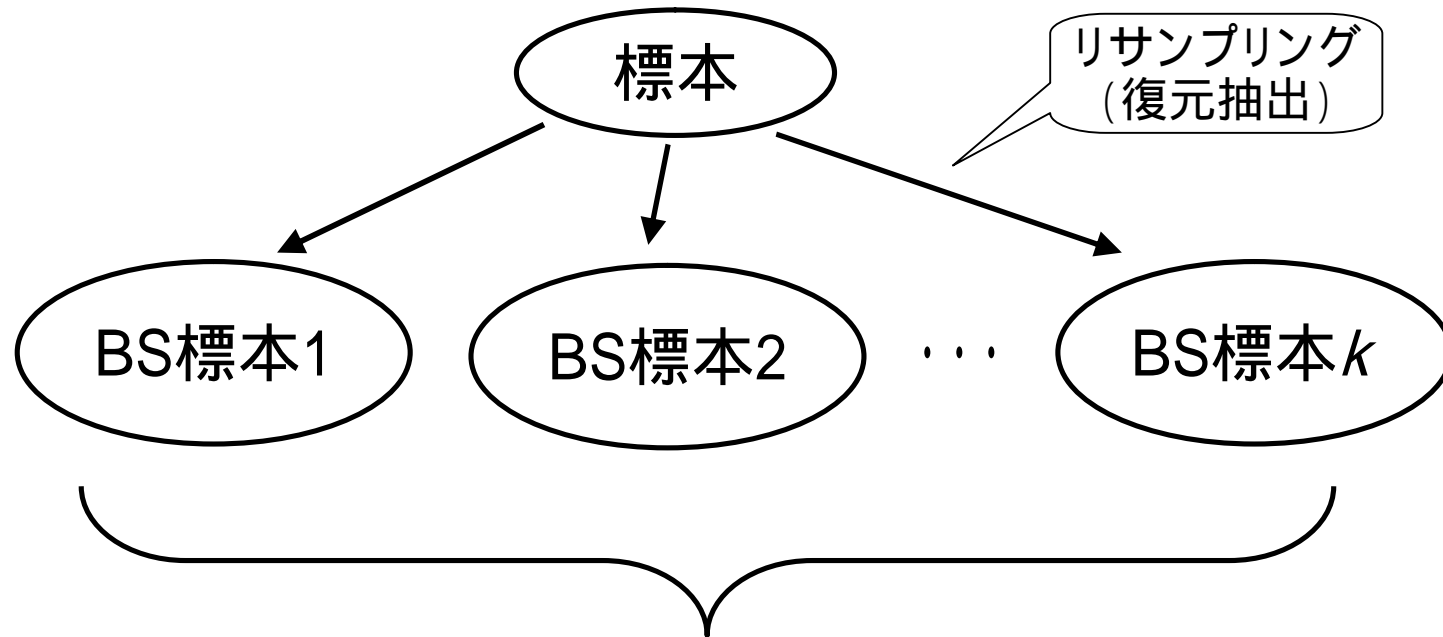


ブートストラップ法を適用することによって…

負荷量の95%信頼区間を求めることができる

ある変数が特定の因子に含まれることの信頼性の向上

ブートストラップ法



k 個の平均値 or 回帰係数が得られる

case-based resampling

元の標本データ					
	X1	X2	X3	...	X _p
No.1	5	4	4	...	3
No.2	4	3	5	...	5
No.3	2	3	3	...	3
⋮					
⋮					
No.i	2	4	5	...	5

ブートストラップ標本					
	X1	X2	X3	...	X _p
	4	3	5	...	5
	5	5	4	...	2
	~~~~~				
	2	4	5	...	5

回答者単位(行単位)でリサンプリングを行う

Davison, Hinkley. (1997). *Bootstrap Method and Their application*. Cambridge.ではこれを **case-based resampling** としている。



# シミュレーションの手順

---

## Step 1

元の標本の大きさと同じ大きさのブートストラップ標本を作成し、

## Step 2

そのブートストラップ標本に対して因子分析（最尤法、プロマックス回転）を実行して因子負荷量を推定する。

## Step 3

これを10000回繰り返し行うことにより、各観測変数に対して10000個の因子負荷量が得られる。

## Step 4

これよりヒストグラムを描き、95% の信頼区間を求めてみる。

# 従来法による結果

結婚のデータセットに対して探索的因子分析を行った結果

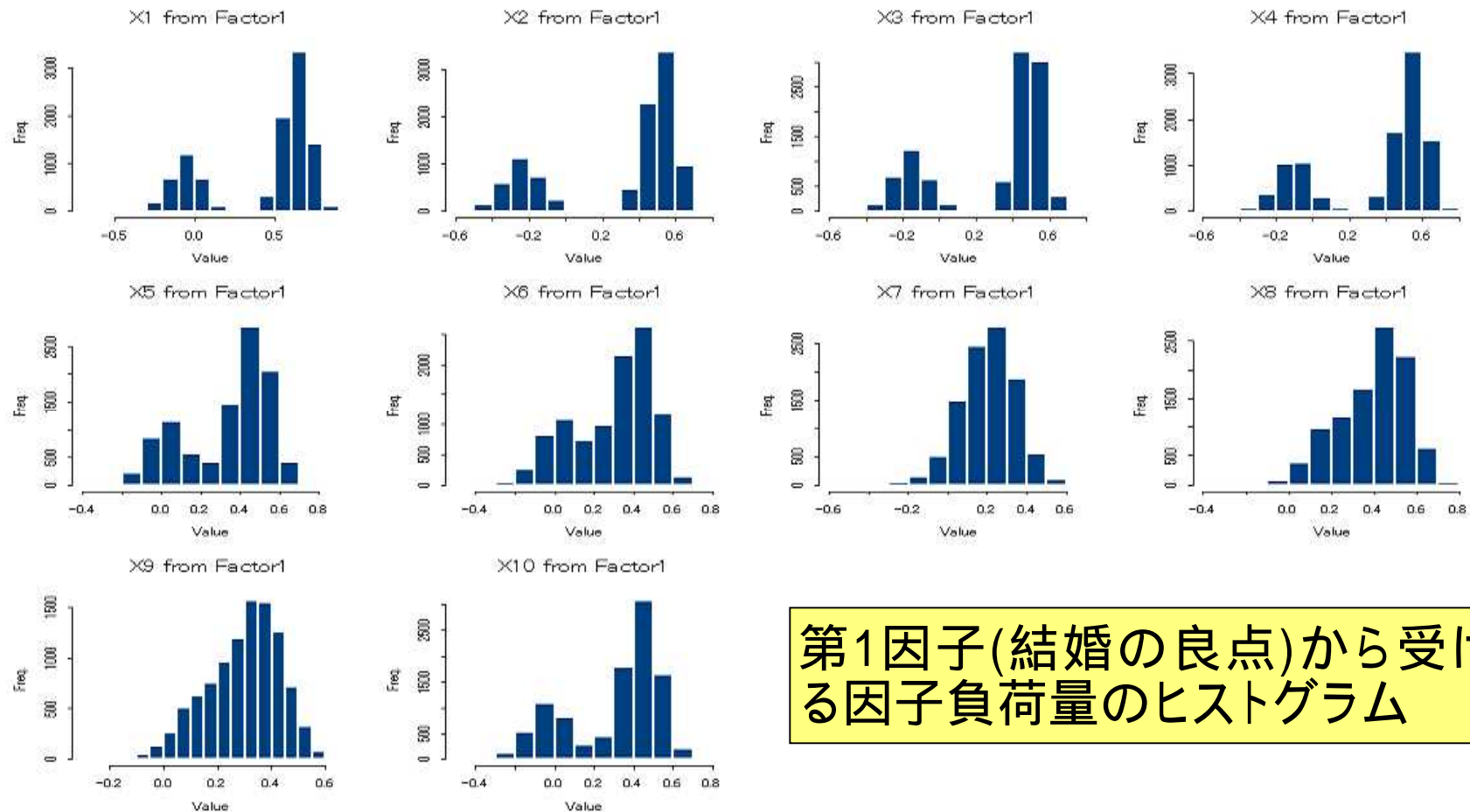
結婚の良い点

項目	第1因子
X1	0.624
X2	0.511
X3	0.488
X4	0.530
X5	0.454
X6	0.407
X7	0.254
X8	0.482
X9	0.358
X10	0.444

項目	第2因子
X11	0.606
X12	0.510
X13	0.481
X14	0.468
X15	0.437
X16	0.431
X17	0.447
X18	0.435

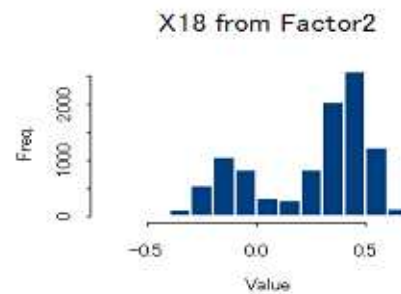
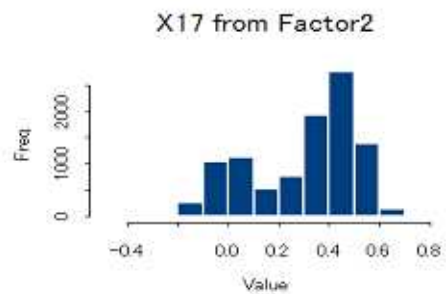
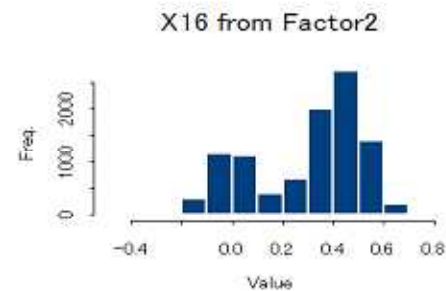
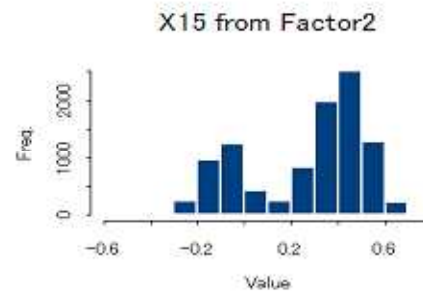
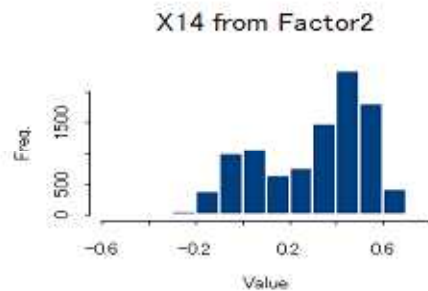
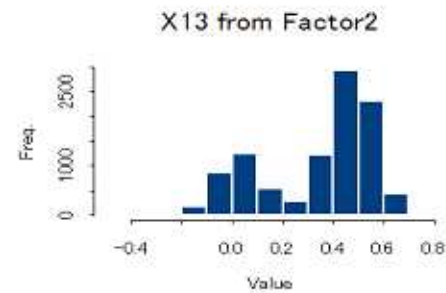
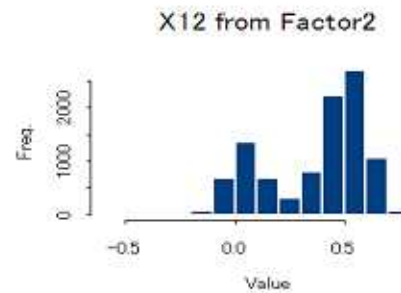
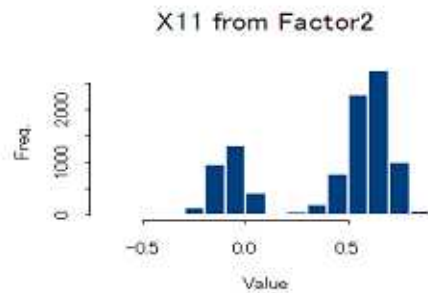
結婚の悪い点

# 探索的因子分析の場合(1)



第1因子(結婚の良点)から受ける因子負荷量のヒストグラム

# 探索的因子分析の場合(2)



第2因子(結婚の悪点)から  
受ける因子負荷量のヒスト  
グラム

# EFAを用いることの問題点

作成されるBS標本の変動によって、因子構造も変化してしまう

あるBS標本では・・・

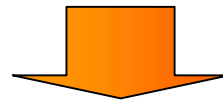
$$F_1 = \{X_1, X_2, \dots, X_{10}\}$$

$$F_2 = \{X_{11}, X_{12}, \dots, X_{18}\}$$

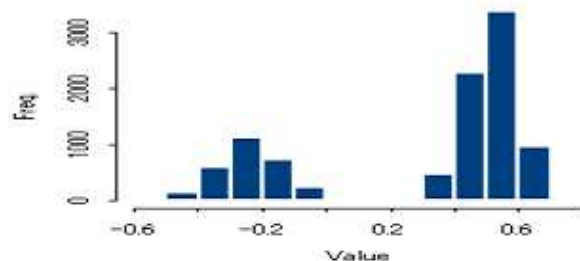
別のあるBS標本では・・・

$$F_1 = \{X_{11}, X_{12}, \dots, X_{18}\}$$

$$F_2 = \{X_1, X_2, \dots, X_{10}\}$$



2つの鐘形の分布が現れることになる



95%信頼区間を求めても  
意味が無い!!

# 検証的因子分析の利用

---

特定の因子構造を仮定して行う(CFA)ことで問題を回避できる

S-PLUSにはCFAを行う関数がないのでRを利用

Sの関数とRの関数は相互利用できる場合が多い

Sユーザなら、Rで定義されているコードを読むこともできる

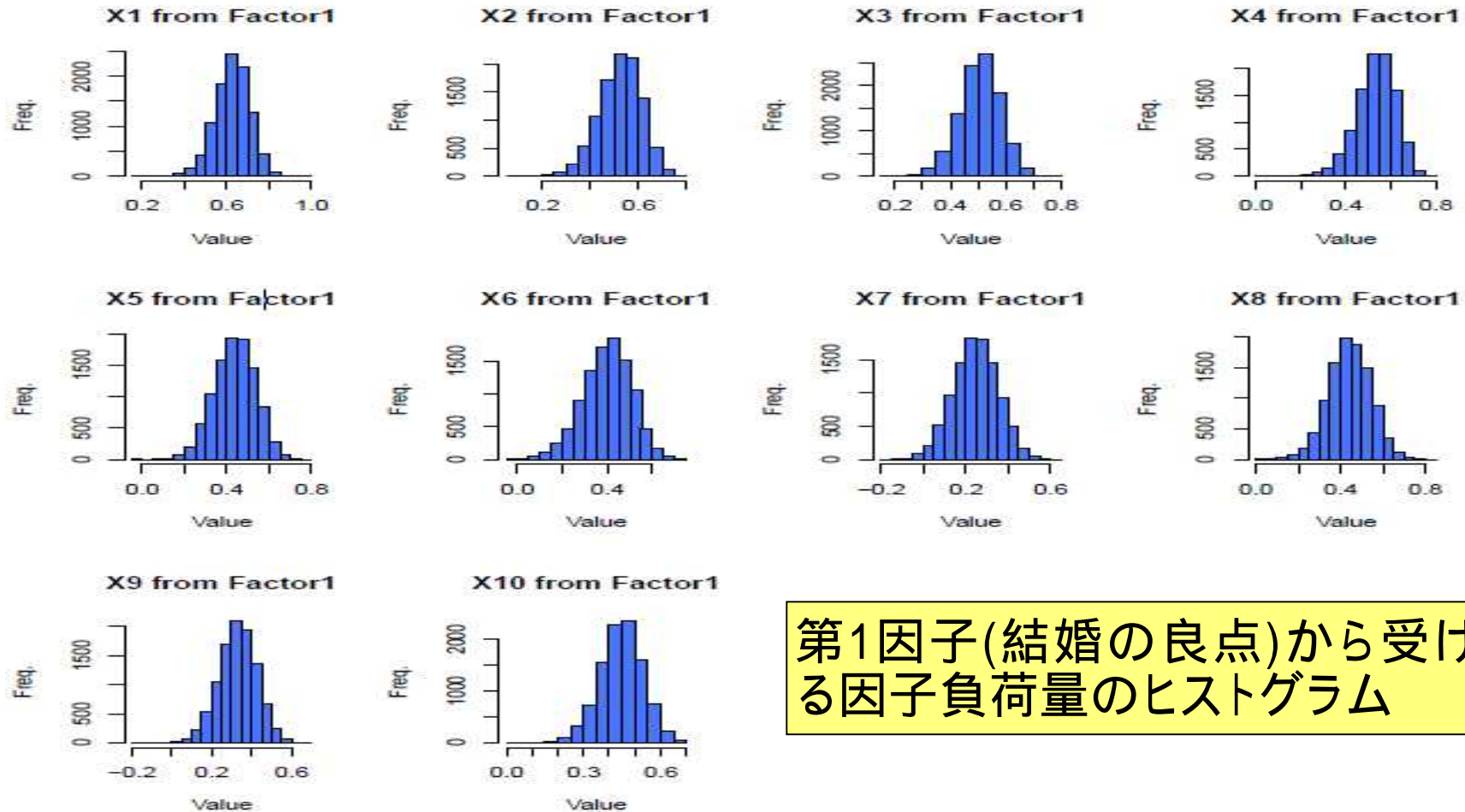
Rでは数々のパッケージが公開されている

Rで定義される`cfa()`をSの`factanal()`の代わりに使用して実行

* 群馬大学の青木繁伸氏が定義しているものを引用

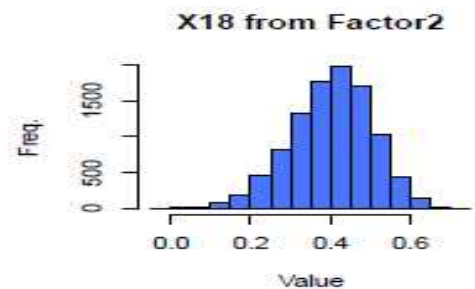
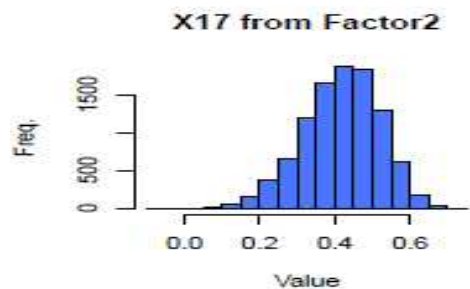
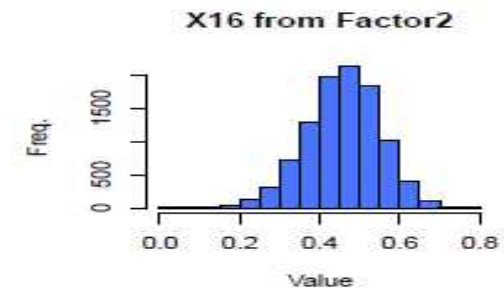
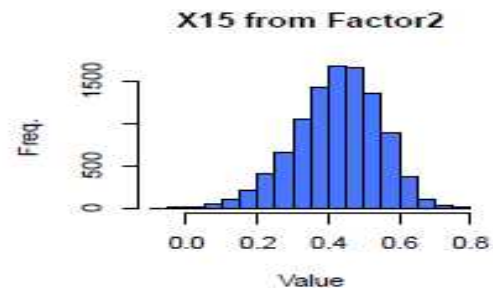
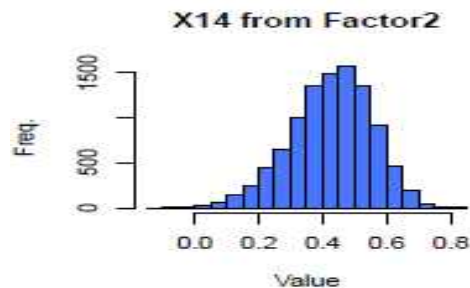
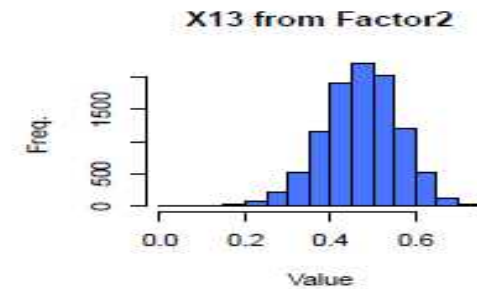
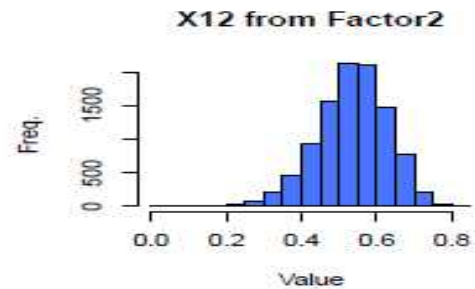
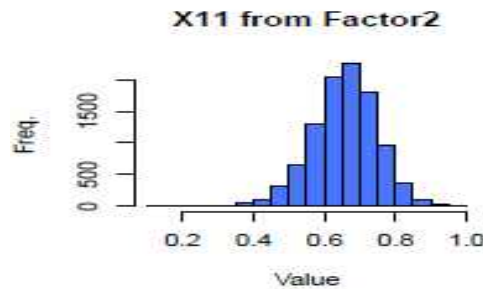
<http://aoki2.si.gunma-u.ac.jp/R/cfa.html>

# 検証的因子分析の場合(1)



第1因子(結婚の良点)から受ける因子負荷量のヒストグラム

# 検証的因子分析の場合(2)

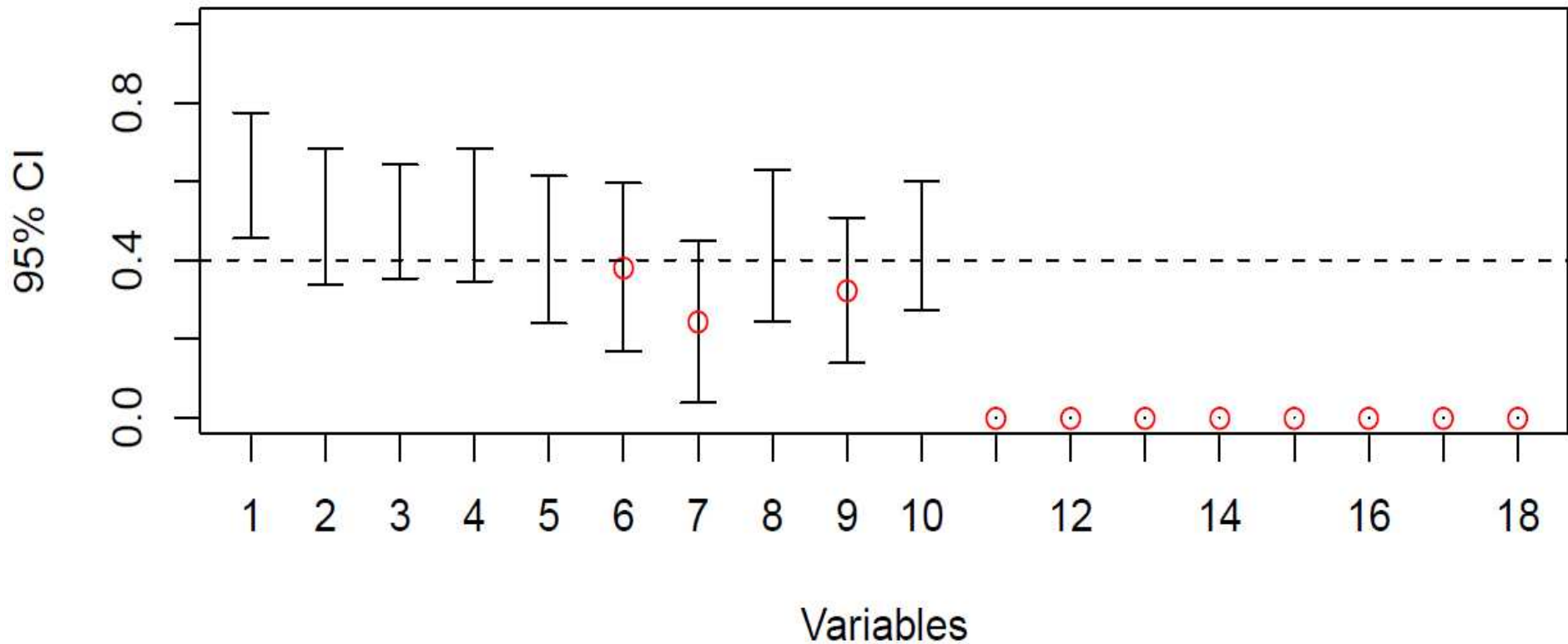


第2因子(結婚の悪点)から  
受ける因子負荷量のヒスト  
グラム



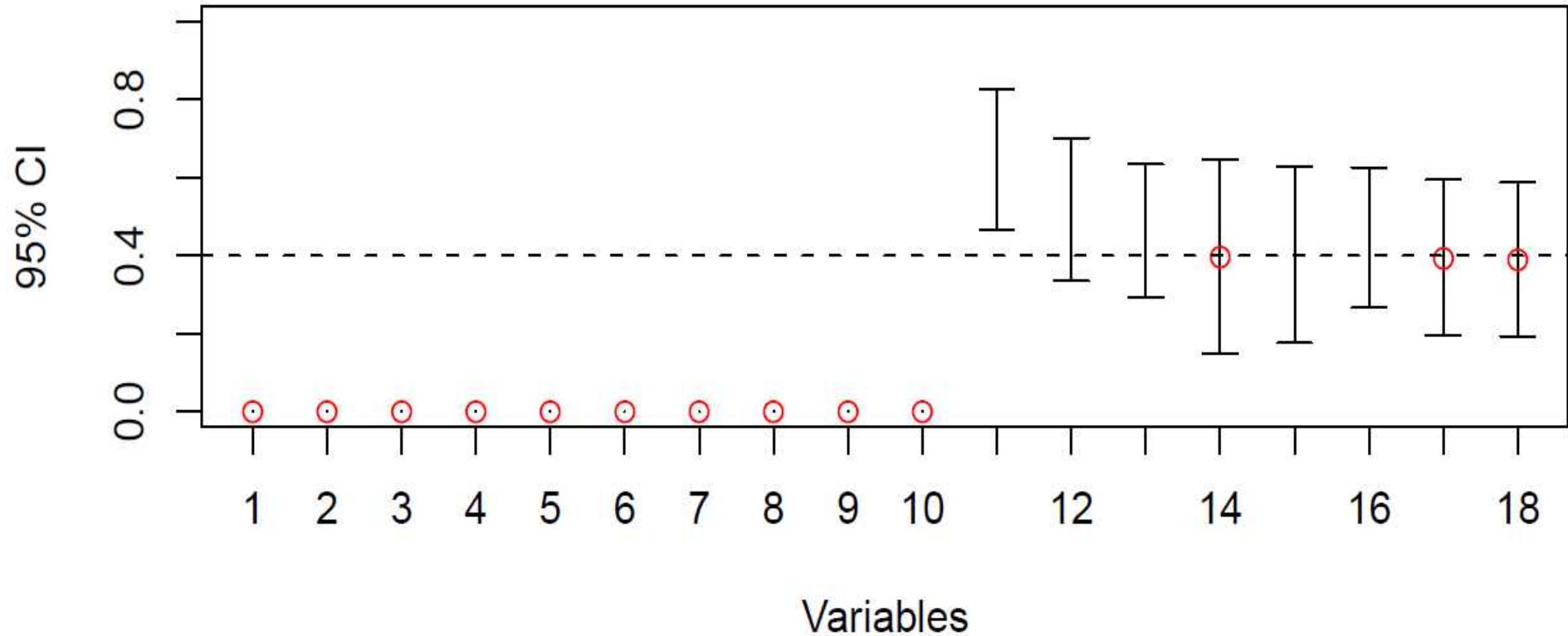
# 因子負荷量の95%信頼区間(1)

第1因子(結婚の良点)から各観測変数への因子負荷量



# 因子負荷量の95%信頼区間(2)

第2因子(結婚の悪点)から各観測変数への因子負荷量



# 従来法とBS法による結果

---

## 従来法による項目分類の結果

第1因子の削除対象 ---> X7, X9

第2因子の削除対象 ---> なし(全ての変数が含まれる)

## ブートストラップ法による項目分類の結果

第1因子の削除対象 ---> X6, X7, X9

第2因子の削除対象 ---> X14, X17, X18

# モデル解析に使用する変数

---

将来的に結婚したいと思うかどうか: *HOPE* (2値データ)

結婚の良い点に関する質問項目の合計点: *GOOD* (連続型)

結婚の悪い点に関する質問項目の合計点: *BAD* (連続型)

---

解析するモデル(2値ロジットモデル):

$$HOPE = GOOD + BAD$$

結婚したいと思うかどうかは、結婚に対する意識によって予測できるか？

# 解析結果の比較

---

従来法で得られた尺度得点を説明変数に用いた場合の結果  
*T.GOOD*によってのみ説明(予測)されるというモデル

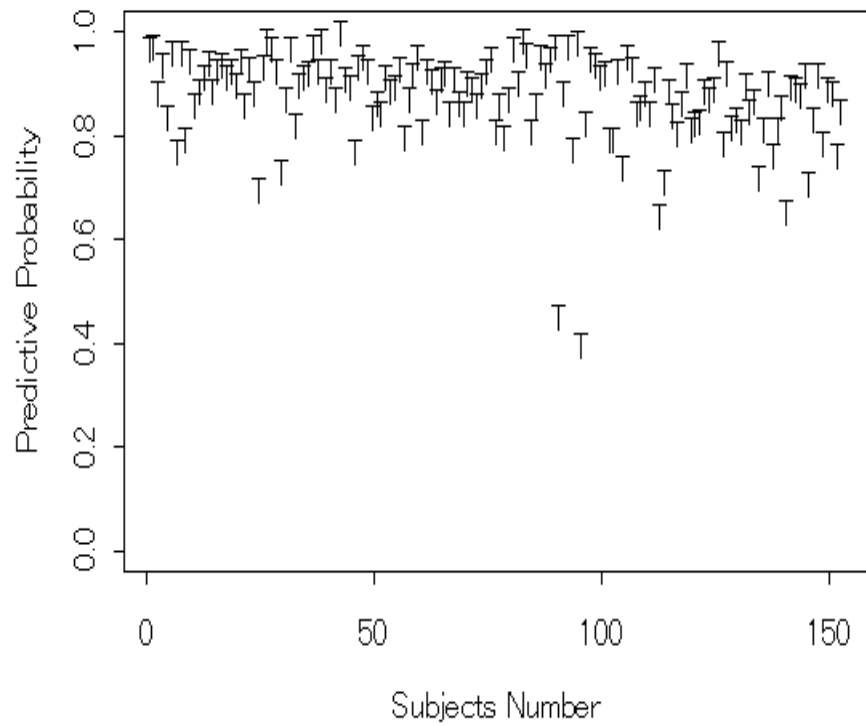
$$HOPE = T.GOOD$$

BS法で得られた尺度得点を説明変数に用いた場合の結果  
*BS.GOOD*と*BS.BAD*の交互作用モデル

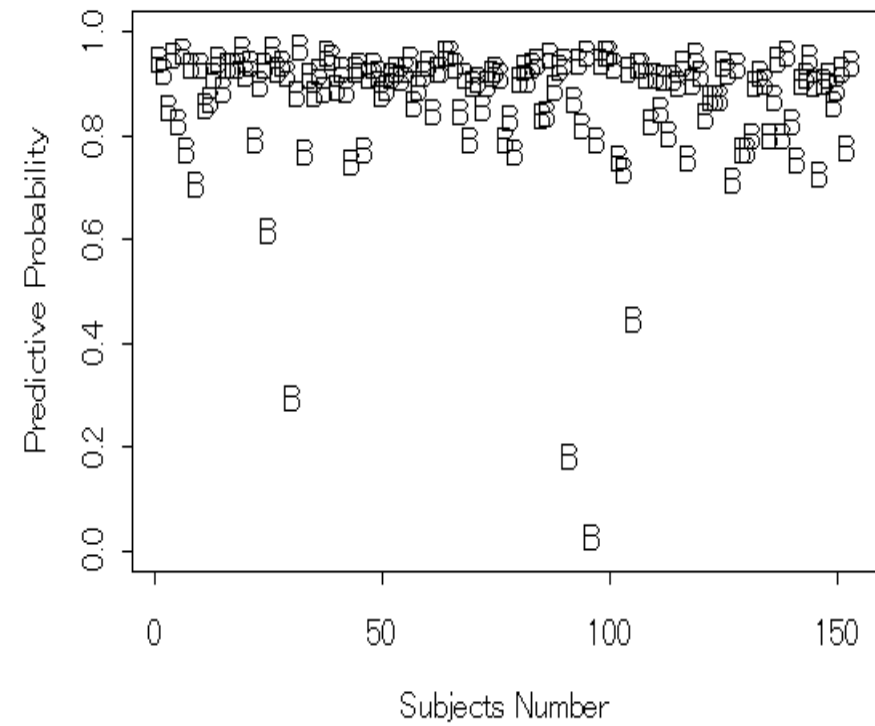
$$HOPE = BS.GOOD * BS.BAD$$

# 予測値のプロット

Usual Methods



Bootstrap Methods



# まとめ

---

cfa()を用いれば、因子負荷量の推定にブートストラップ法を適用できる

95%信頼区間を求めることで、負荷量が実質的に0であるかどうかを判断できる(=負荷量が0であるという仮説に対する検定の役割)

従来法とBS法とで計算される尺度得点によって、その後のモデル解析の結果も異なることがある

今回の事例では、BS法による尺度得点を用いたモデルの方が、より正確に予測することができた

# 問題点と考察

---

従来法とBS法とで項目分類の結果が異なる場合もある

積極的解釈：従来法の結果をより信頼あるものと主張できる

消極的解釈：所詮トートロジーに過ぎないのではないか

削除の対象となる項目に違いがあっても、結果はさほど変わらない

計算時間もかかるので冗長な作業といえる

ブートストラップ標本の大きさはどれ位が適切か

因子分析の場合、10000回の繰り返しより多くが必要かもしれない



# おわりに S-PLUSの活用

---

シミュレーションによる理論の理解

S言語を使えば、初心者でもプログラミングできる

数式による理解が難しい場合でも視覚的に納得できる

S-PLUSのGUI機能の活用

コマンド操作に慣れていない人でも入門しやすい

複雑なグラフ作成においては、コマンド操作よりも簡単である

Rとの互換性

Rに実装されている関数も流用することができる

感謝

ご清聴ありがとうございました

目白大学 心理学研究科

片所 強（かたしよ つよし）

E-mail: [katasho244@nifty.com](mailto:katasho244@nifty.com)