

(2021-JA-Networking)

PLS回帰

九州大学 岸本 淳司

JAPAN 2021



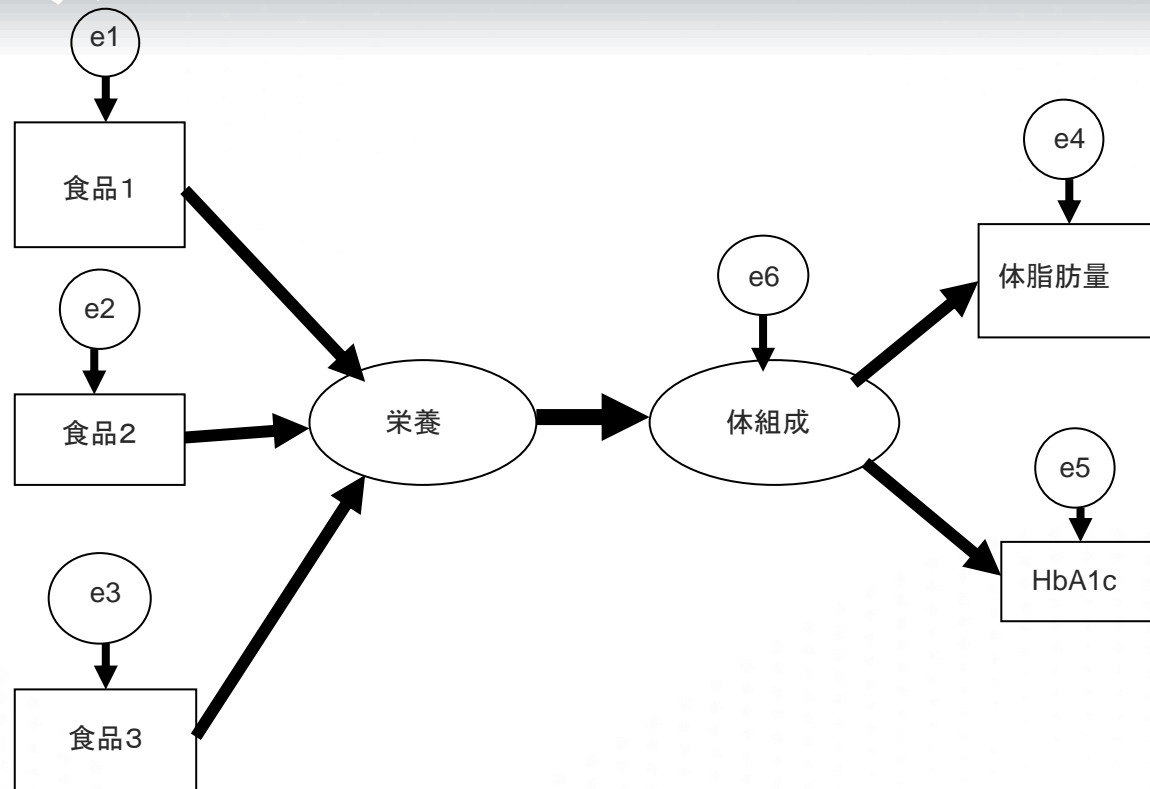
問題: DIET Studyでのサブ解析

- 日本人2型糖尿病患者を対象としたダパグリフロジン (SGLT2阻害薬)が投与された際の体重減少時の体組成と食行動に及ぼす影響を検討する2年間の臨床試験 (DIET study)
- <サブ仮説>患者のベースラインの食行動によってSGLT2阻害薬の効果(体脂肪量減少やHbA1c減少など)の程度が異なるのではないか

データ構造

- 応答変数
 - 2年時点の体脂肪量
 - HbA1cのベースラインからの変化量
- 説明変数
 - 自記式質問票BDHQ(58品目→33食品群)
 - 治療開始前・開始後の食品および栄養素摂取量

PLSモデル



PLSモデルの特徴

- 構造方程式モデルの1つとして教育されることが多い
- 原因系変数も結果系変数も複数ある回帰モデル
- 通常の回帰分析が困難な場合でも使える
 - 説明変数の数が症例数より多い
 - 説明変数間に高い相関がある
- 多因子モデルも扱える

分析 → 多変量 → PLS回帰 (例題: Crime.jmp)

Crime - JMP Pro

ファイル(F) 編集(E) テーブル(T) 行(R) 列(C) 実験計画(DOE)(D) 分析(A) グラフ(G) ツール(O) アドイン(N) 表示(V) ウィンドウ(W) ヘルプ(H)

Crime

ロックされたファイル C:*
ノート Crime Rates per

列(8/0)

州

州	夜盗	窃盗	車の窃盗	
1 ALABAMA	1135.5	1881.9	280.7	
2 ALASKA	1331.7	3369.8	753.3	
3 ARIZONA	2346.1	4467.4	439.5	
4 ARKANSAS	972.6	1862.1	183.4	
5 CALIFORNIA	2139.4	3499.8	663.5	
6 COLORADO				
7 CONNECTICUT				
8 DELAWARE				
9 FLORIDA				
10 GEORGIA				
11 HAWAII				
12 IDAHO	5.5	19.4	39.6	172.5
13 ILLINOIS	9.9	21.8	211.3	209.0
14 INDIANA	7.4	26.5	123.2	153.5
15 IOWA	2.3	10.6	41.2	89.8
16				

行

すべての行 50
選択されている行 0
除外されている行 0
表示しない行 0
評価が完了しました。

分析(A) のメニュー:

- 一変量の分布
- 二変量の関係
- 表の作成
- テキストエクスプローラ
- モデルのあてはめ
 - 予測モデル
 - 発展的なモデル
 - スクリーニング
- 多変量
 - クラスタ分析
 - 品質と工程
 - 信頼性/生存時間分析
 - 消費者調査
- 多変量の相関
- 主成分分析
- 判別分析
- PLS回帰
- 多重対応分析
- 構造方程式モデル
- 因子分析
- 多次元尺度構成
- 項目分析

PLS回帰ダイアログ

Crime - JMP Pro

PLS回帰 - JMP Pro

説明変数の線形結合である因子に基づいて、線形モデルをあてはめる。

列の選択

8列

- 州
- 殺人
- 強姦
- 強盗
- 暴行
- 夜盗
- 窃盗
- 車の窃盗

アクション

目的変数

- 夜盗
- 窃盗
- 車の窃盗

オプション(数値)

説明変数

- 殺人
- 強姦
- 強盗
- 暴行

度数

検証

By

オプション

中心化

尺度化

欠測データの補完

OK

キャンセル

削除

前回の設定

ヘルプ

	窃盗	車の窃盗
5.5	1881.9	280.7
1.7	3369.8	753.3
6.1	4467.4	439.5
2.6	1862.1	183.4
9.4	3499.8	663.5
5.2	3903.2	477.1
6.0	2620.7	593.2
2.6	3678.4	467.0
9.9	3840.5	351.4
1.1	2170.2	297.9
1.5	3920.4	489.4
0.8	2599.6	237.6
5.0	2828.5	528.6
0	2498.7	377.4
0	2685.1	219.9

評価が完了しました。

「モデルのあてはめ」からも実行可能

モデルのあてはめ - JMP Pro

▼ モデルの指定

列の選択

8列

- 州
- 殺人
- 強姦
- 強盗
- 暴行
- 夜盗
- 窃盗
- 車の窃盗

役割変数の選択

Y

- 夜盗
- 窃盗
- 車の窃盗

重み オプション(数値)

度数 オプション(数値)

検証 オプション(数値)

By オプション

手法: 標準最小2乗

強調点:

- 標準最小2乗
- ステップワイズ法
- 一般化回帰
- 混合モデル
- MANOVA
- 対数線形-分散
- 名義ロジスティック
- 順序ロジスティック
- 比例ハザード
- 生存時間(パラメトリック)
- 一般化線形モデル
- PLS回帰
- 応答のスクリーニング

個別化

ヘルプ

前回の設

削除

モデル効果の構成

追加 殺人

交差 強姦

枝分かれ 強盗

マクロ ▼ 暴行

次数 2

属性 ▼

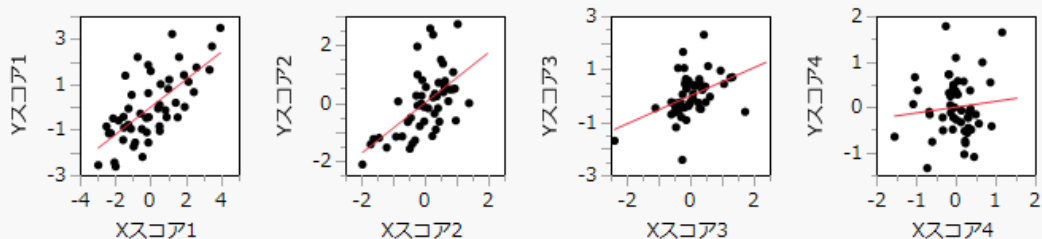
変換 ▼

切片なし

にする

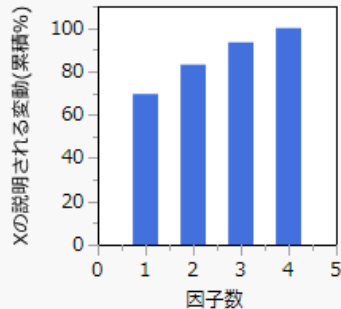
基本出力(多因子解)

X-Yスコアプロット

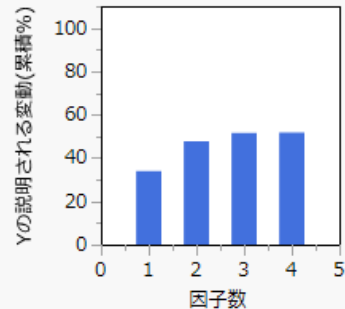


説明される変動(%)

因子数	X効果	20 40 60 80	累積X
1	69.4898		69.490
2	13.5819		83.072
3	10.2623		93.334
4	6.6660		100.000



Y応答	20 40 60 80	累積Y
34.0204		34.0204
13.7534		47.7739
3.8575		51.6314
0.1446		51.7760

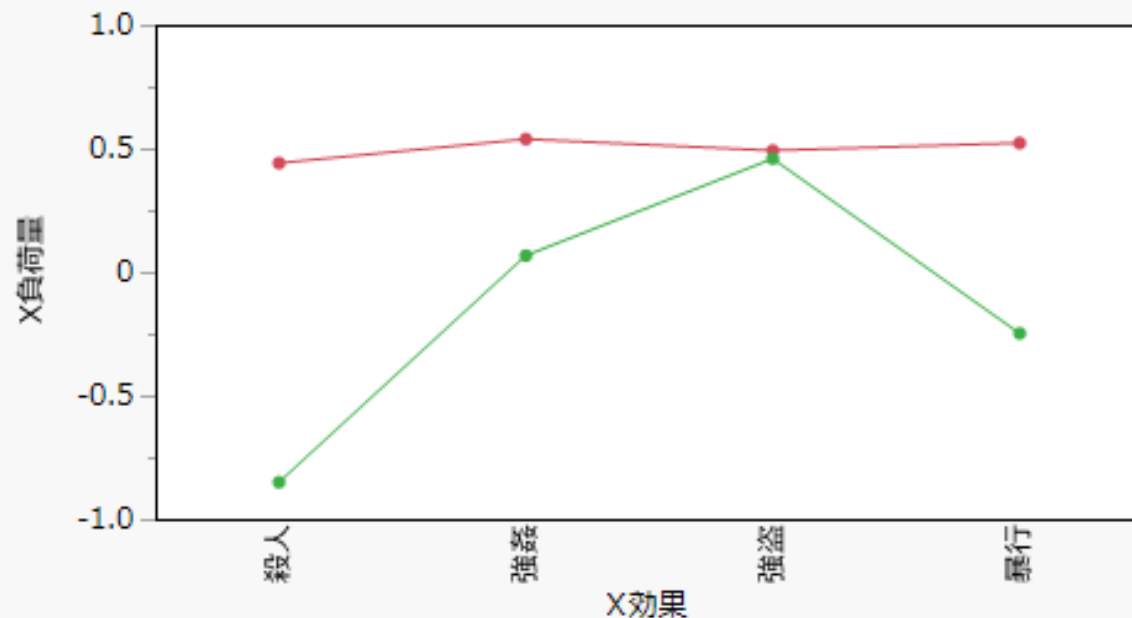


X側負荷量プロット(2因子解)

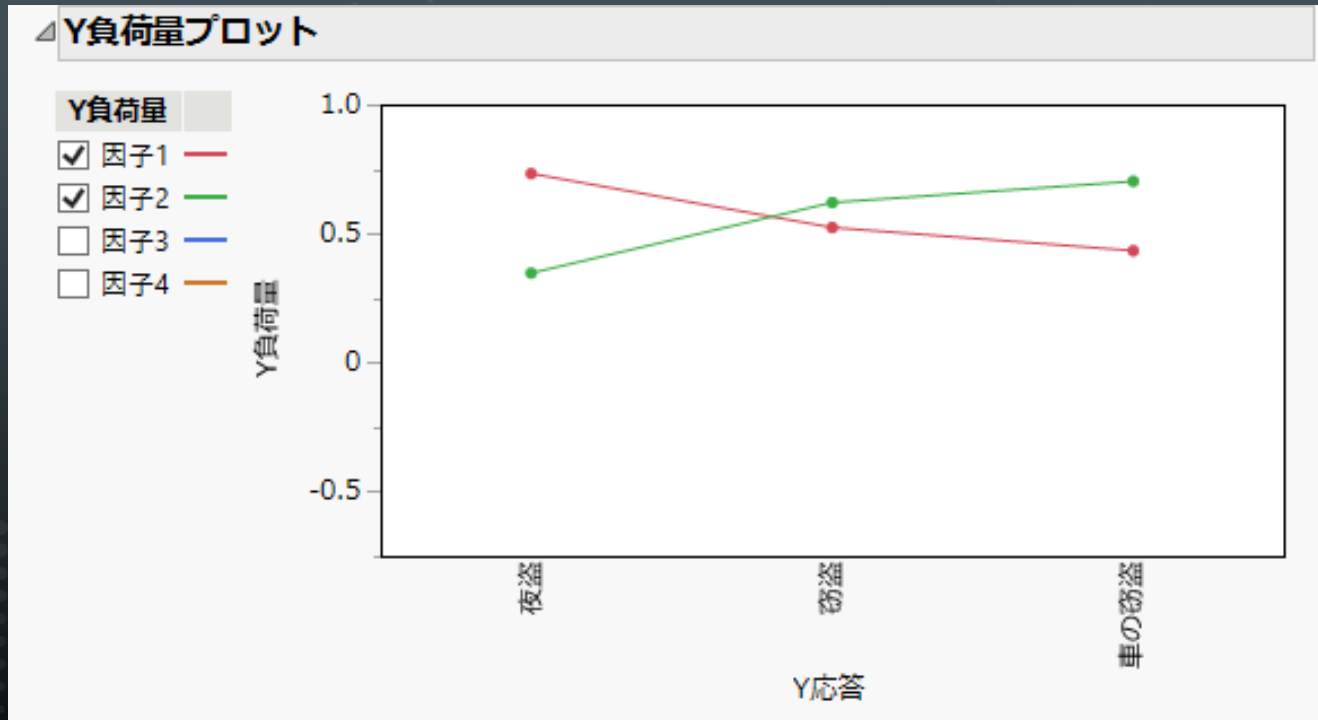
△ X負荷量プロット

X負荷量

- 因子1
- 因子2
- 因子3
- 因子4



Y側負荷量プロット(2因子解)



豊富なデータセット出力

The screenshot displays the SAS JMP interface for PLS regression. The left pane shows the 'PLS 回帰' (PLS Regression) window with the following settings:

- モデルの設定** (Model Settings): NIPALS 高速
- モデル比較の要約** (Model Comparison Summary):

手法	特異値分解	行数	因子数	Xの説明され
NIPALS	高速	50	4	
- [NIPALS]: [4]因子, 特異値分解=[高速]** (Selected Model):
 - 変動プロット
 - 変数重要度のプロット
 - 変数重要度 vs 係数 プロット
 - VIP 関連の設定
 - 係数プロット
 - 負荷量プロット
 - 負荷量散布図行列
 - 負荷量の相関図
 - X-Yスコアプロット
 - スコア散布図行列
 - 距離プロット
 - T2乗プロット
 - 診断プロット
 - プロフィール
 - スペクトルプロフィール
 - 列の保存

The right pane shows the '予測式の保存' (Save Prediction Equations) dialog box with the following options:

- 予測式の保存
 - 予測値をXスコアの計算式として保存
 - Xの予測値をXスコアの計算式として保存
 - 予測値の標準誤差の計算式を保存
 - 平均の信頼限界の計算式を保存
 - 個別の信頼限界の計算式を保存
 - スコアの計算式を保存
- Y予測値の保存
 - Y残差の保存
 - X予測値の保存
 - X残差の保存
 - Xの説明される変動を保存
 - Yの説明される変動を保存
 - スコアの保存
 - 負荷量の保存
 - 標準化したスコアを保存
 - 標準化した負荷量を保存
 - T2乗の保存
 - T2乗をXスコアの計算式として保存
 - 距離の保存
 - 距離をXスコアの計算式として保存
 - X重みの保存
- 予測式を発行
 - スコアの計算式を発行

まとめ

- JMPによるPLS回帰は、非常に簡単な指定で実行できる
- 多因子解が自然に得られる
- 結果出力／データセット出力が工夫されて使いやすい

JAPAN
**DISCOVERY
SUMMIT**
EXPLORING DATA
INSPIRING INNOVATION

YOUR TURN

