

Plotの理解

筑波大学医学医療系精神医学
根本清貴

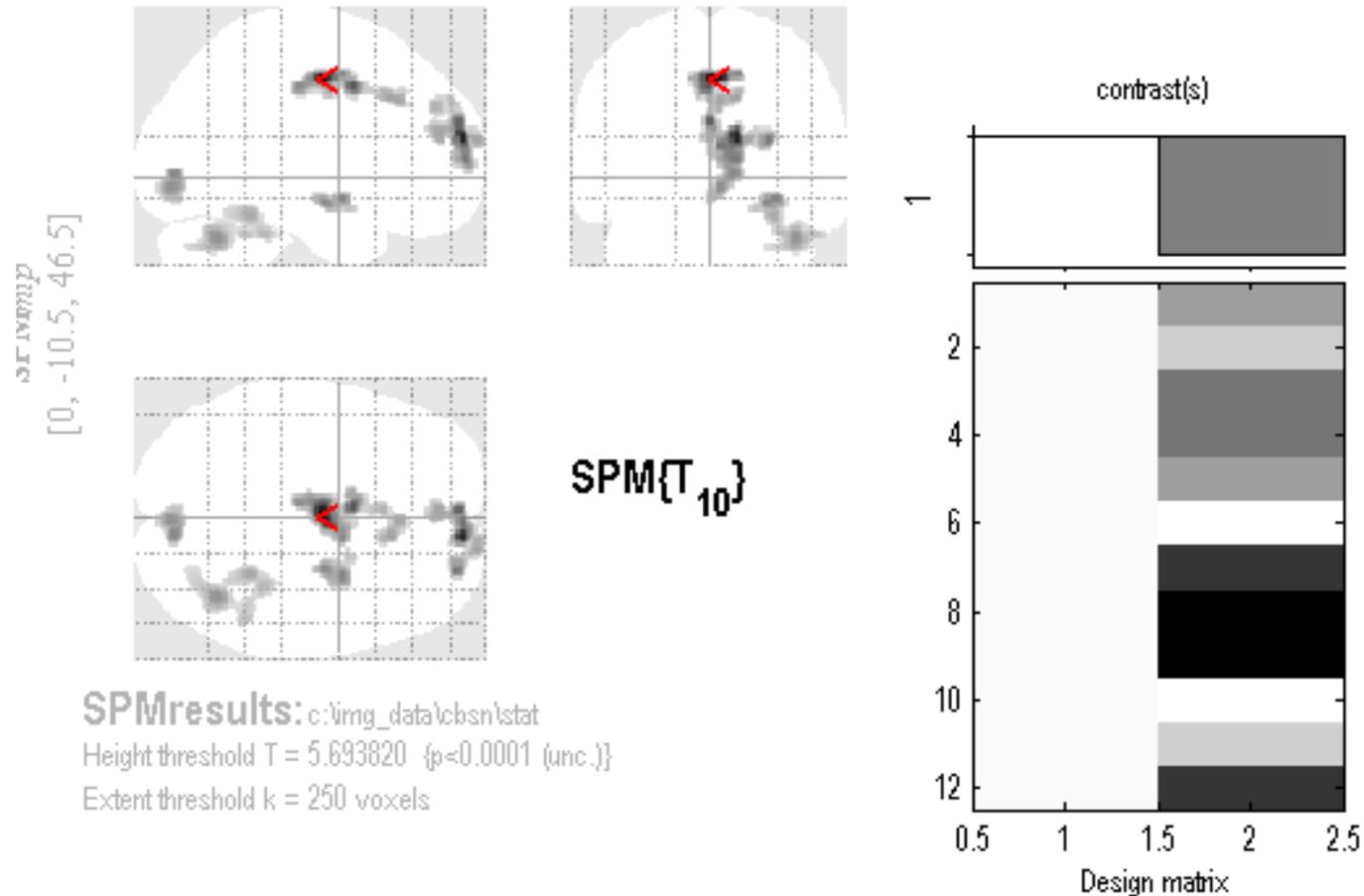
一般線形モデルとは

- 説明変数の線形結合に残差項を加えて目的変数を説明するモデル
 - 目的変数を以下の式で説明しようとする

$$Y = X \times B + E$$

- 年齢と脳容積の相関の場合
 - Y: 目的変数 = 画像データ (灰白質容積)
 - X: 計画行列 (年齢、その他)
 - B: まだわからないパラメータ
(回帰直線の傾きと切片)
 - E: 回帰直線では説明しきれないゆらぎ

相関解析(年齢と負の相関を示す領域)

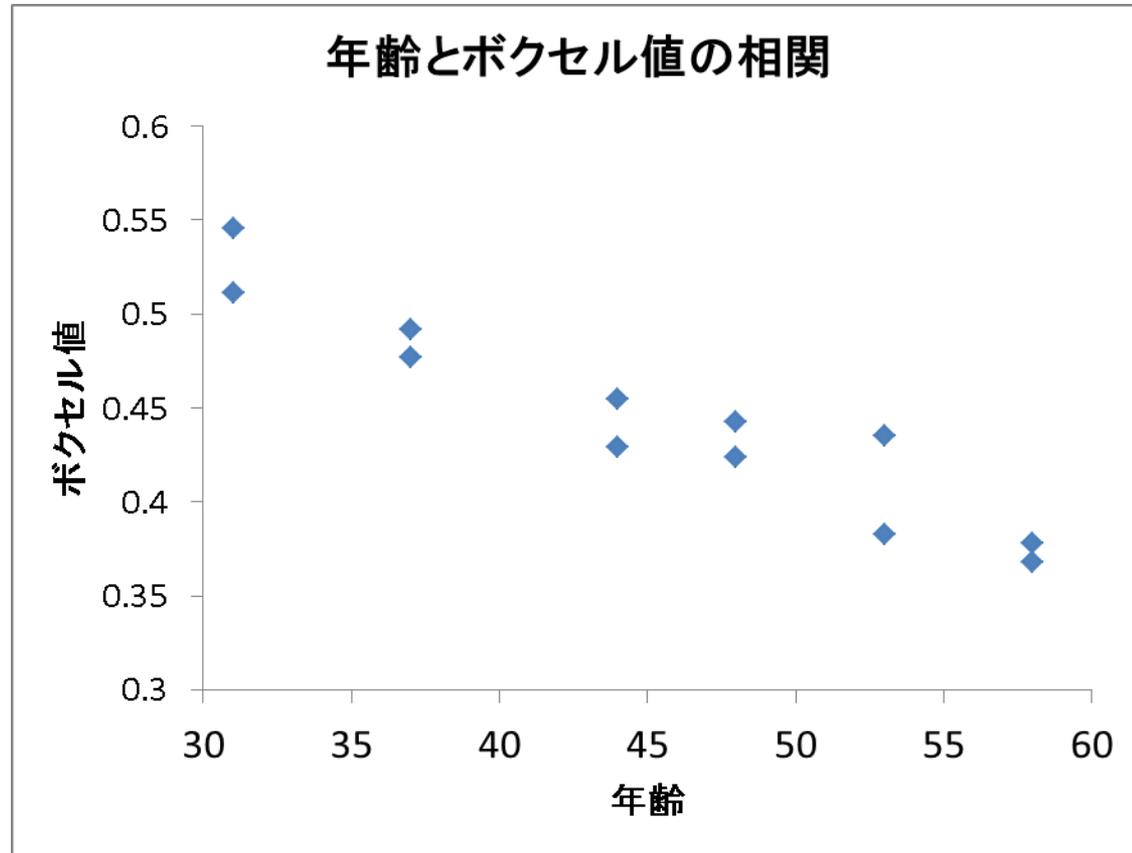


Statistics: p-values adjusted for search volume

| set-level | | cluster-level | | | | peak-level | | | | | mm mm mm | | |
|-----------|---|----------------------|----------------------|----------------|--------------------|----------------------|----------------------|-------|-------------------|--------------------|----------|-----|----|
| p | c | p _{FWE-cor} | q _{FDR-cor} | k _E | p _{uncor} | p _{FWE-cor} | q _{FDR-cor} | T | (Z _z) | p _{uncor} | | | |
| 0.000 | 5 | 0.000 | 0.000 | 898 | 0.000 | 0.034 | 0.574 | 12.00 | 5.13 | 0.000 | 0 | -10 | 46 |
| | | | | | | 0.150 | 0.574 | | | | | | |
| | | | | | | 0.336 | 0.574 | 8.12 | 4.41 | 0.000 | 11 | 0 | 48 |

ある座標における ボクセル値と年齢の散布図

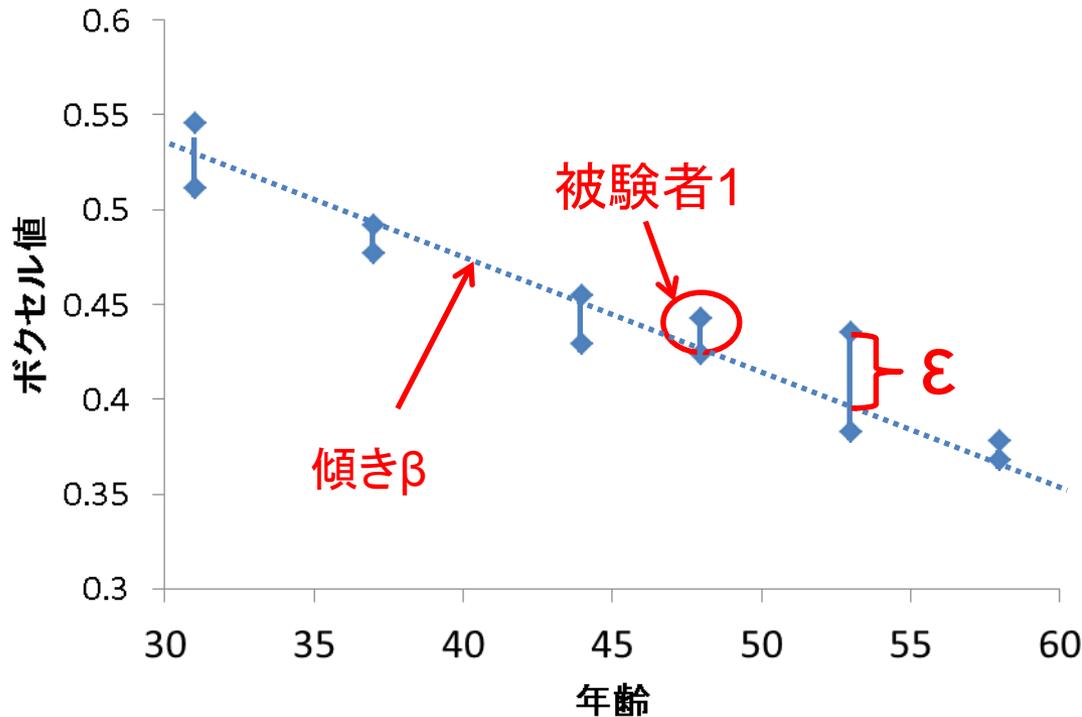
| 被験者 | ボクセル値 (0 -10 46) | 年齢 |
|-----|---------------------|----|
| 1 | 0.42 | 48 |
| 2 | 0.44 | 53 |
| 3 | 0.43 | 44 |
| 4 | 0.45 | 44 |
| 5 | 0.44 | 48 |
| 6 | 0.38 | 58 |
| 7 | 0.49 | 37 |
| 8 | 0.51 | 31 |
| 9 | 0.55 | 31 |
| 10 | 0.37 | 58 |
| 11 | 0.38 | 53 |
| 12 | 0.48 | 37 |



仮説：年齢とボクセル値は負の相関を示すのではないか？

年齢とボクセル値の関係が直線で表示されると仮定

年齢とボクセル値の相関



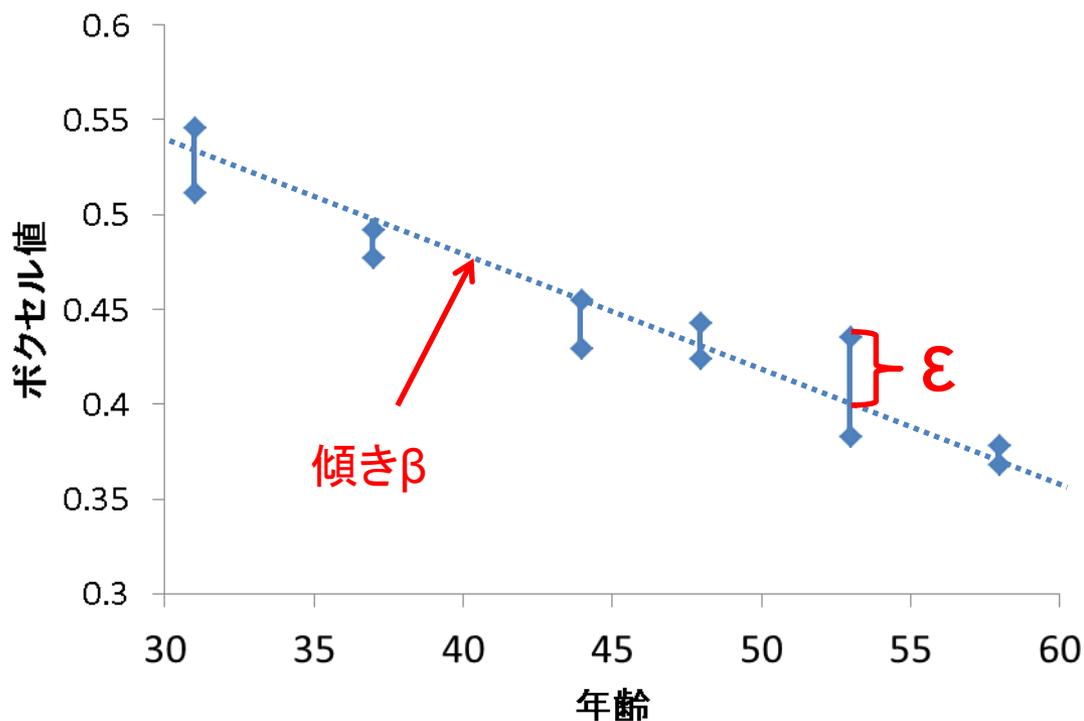
$$Y_j = c + \beta \times x_j + \varepsilon_j$$

- Y : ボクセル値
- j : 被験者番号
- β : 直線の傾き
- x : 年齢
- c : 切片 (年齢が0の時のボクセル値)
- ε : ゆらぎ (残差)

被験者1ならば、ボクセル値0.42、年齢48歳なので、
 $0.42 = c + \beta \times 48 + \varepsilon_1$ と表現できる。

もっともフィットする直線は？

年齢とボクセル値の相関



- もっともフィットする直線は、残差 ε を最小にする直線
- 残差の二乗の和を求め、その値を最小にする β を見つける(最小二乗法)

$$\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2$$

\hat{B} から残差行列Eを求める

$$Y = X \times B + E$$

- Bの推定値 \hat{B}

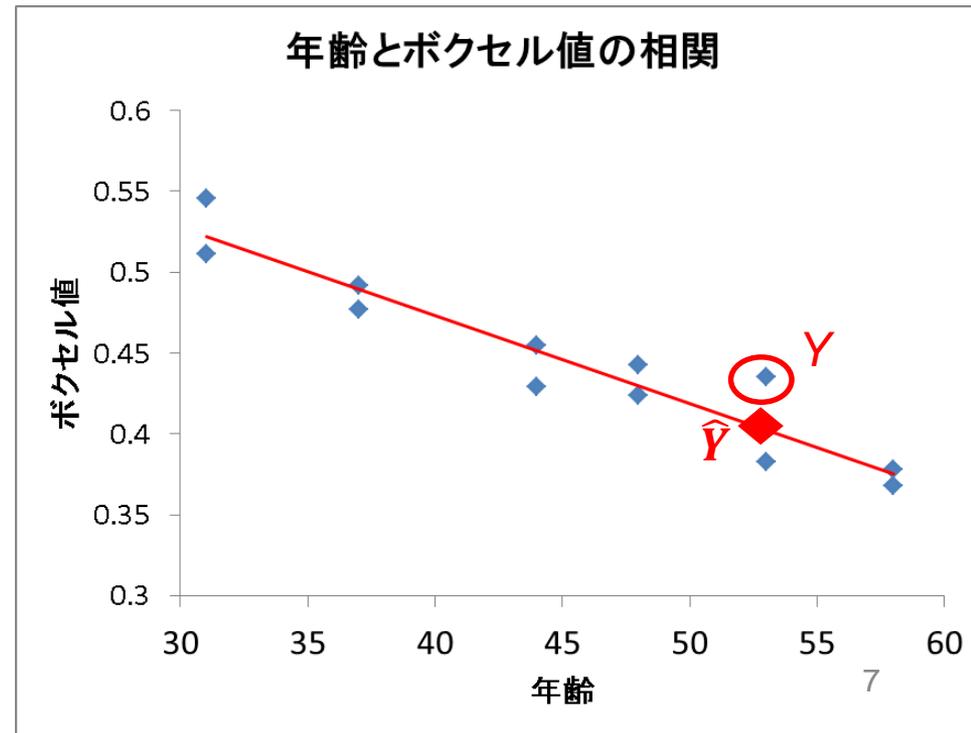
$$\hat{B} = X^+ \times Y$$

- Yの予測値 \hat{Y} は直線上にある。

$$\hat{Y} = X \times \hat{B}$$

- 残差Eの推定値 \hat{E} は実データYと予測値 \hat{Y} の差分。

$$\hat{E} = Y - \hat{Y}$$



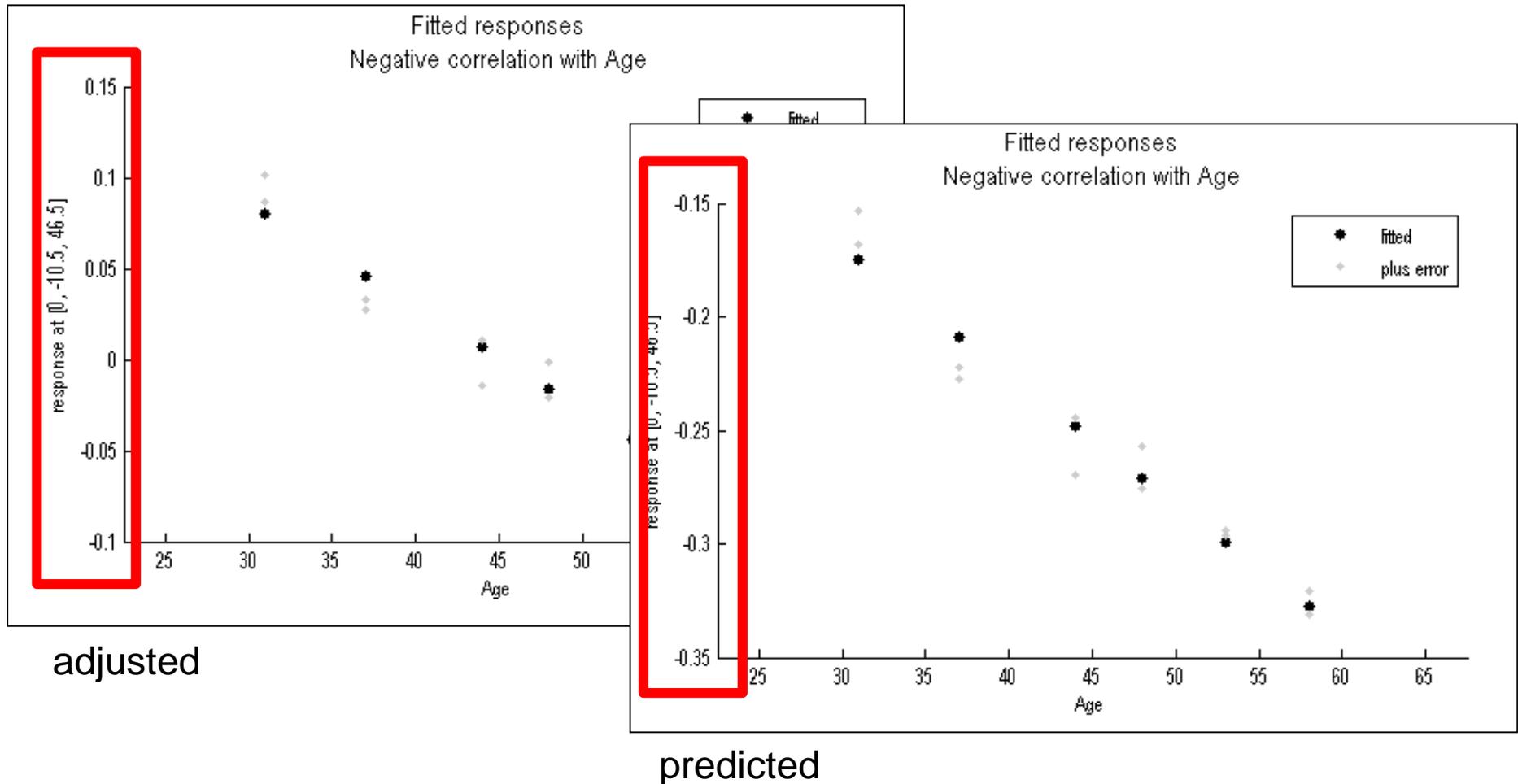
現実：SPMのPlotではこうならない

- SPMの結果からPlotとすると、
 - Predicted
 - Adjustedの2つが聞かれ、どちらかを選んでグラフを作成するものの、Y軸のスケーリングが元データと同じにならない
- そもそも、Predicted, Adjustedの違いは何なのか？

Results → Plot

The image shows a sequence of SPM8 software windows. The top window is titled "SPM8 (Kiyotaka): SPM[T]: Results" and contains a "Design Contrasts" menu. Below the menu are three columns of buttons: "p-values" (whole brain, current cluster, small volume), "Multivariate" (eigenvariate, multivariate Bayes, BMS), and "Display" (CVA, plot, overlays..., save...). The "plot" button in the "Display" column is highlighted with a red box. Below this window is a smaller, identical window, also with the "plot" button highlighted. The bottom window is titled "SPM8 (Kiyotaka): SPM[T]: Results" and shows "Fitted responses" for "Negative correlation with Age". It includes a dropdown menu "Which explanatory variable?..." with "Age" selected. Below the dropdown are input fields for "predicted or adjusted response?" and "predicted".

Predicted vs Adjusted



- 何が違うか？ → Y座標の値が違う
- この違いはどこから？

Plotに関する記載

コマンド ウィンドウ

```
>>  
>> help spm_graph  
Graphical display of adjusted data  
FORMAT [Y y beta Bcov] = spm_graph(xSPM,SPM,hReg)  
  
xSPM - structure containing SPM, distributional & filtering details  
      about the excursion set  
SPM - structure containing generic details about the analysis  
hReg - handle of MIP register or [x y z] coordinates  
  
Y - fitted data for the selected voxel  
y - adjusted data for the selected voxel  
beta - parameter estimates (ML or MAP)  
Bcov - Covariance of parameter estimates (ML or conditional)  
  
See spm_getSPM for details.
```

- Y: fitted
- y: adjusted

Plotに関する混乱

- Adjustedが2つのコンテキストで使われている
 - PredictedとAdjusted
 - FittedとAdjusted
- この混乱を避けるため、以下のように表現
 - Adjusted(y)
 - Adjusted(Y)
 - Predicted(y)
 - Predicted(Y)

Plotの値を検討

- あるボクセルのボクセル値VV(生値)を抽出
- Plotで得られる6つの値を抽出
 - Adjusted(y)
 - Adjusted(Y)
 - Predicted(y)
 - Predicted(Y)
 - beta
 - c
 - β

$$beta = B = \begin{pmatrix} c \\ \beta \end{pmatrix}$$

ボクセル値 (VV) を取り出す

The image shows a sequence of three screenshots from the SPM8 software interface, illustrating the steps to extract voxel values (VV) from a region of interest (ROI).

Top Screenshot: The 'VOI time-series extraction' dialog box is open. The 'Multivariate' tab is selected, and the 'eigenvariate' button is highlighted with a red box. Other buttons include 'whole brain', 'CVA', 'current cluster', 'multivariate Bayes', 'small volume', 'BMS', 'p-value', 'plot', 'overlays...', and 'save...'. The 'Design' tab shows 'VOI time-series extraction: at [0 -10 46]' and 'name of region' set to 'VOI'. The 'VOI definition...' section has 'sphere', 'box', 'cluster', and 'mask' buttons, with 'sphere' and 'box' highlighted by a red box.

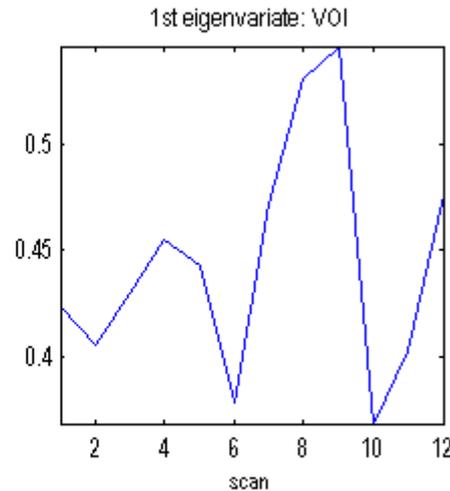
Middle Screenshot: A smaller version of the 'VOI time-series extraction' dialog box is shown, also with 'eigenvariate' and 'sphere/box' buttons highlighted. It includes a 'co-ordinates' section with 'x = 0.00', 'y = -10.50', 'z = 46.50', and 'statistic' set to '12.00'. There are 'clear', 'exit', and '?' buttons.

Bottom Screenshot: The 'VOI time-series extraction' dialog box is shown with the 'sphere' button selected. The 'sphere radius (mm)' field is set to '0'. The 'VOI time-series extraction' section is the same as in the previous screenshots.

単一ボクセルならばsphereでもboxでも可。Boxの1.5はDARTEL後のボクセルが1.5 x 1.5 x 1.5のため。Sphereも1でもよいが、0のままでそのボクセルの値を抽出できる

変数Yにボクセル値が出力される

VOI saved as C:\img_data\cbsn\stat\VOI_VOI.mat



1 voxels in VOI at [0 -10 46]
Variance: 100.00%

Y =

```
0.423709571361539
0.405094176530835
0.429436743259427
0.454467822449109
0.442385911941525
0.378248810768125
0.471460431814191
0.531101763248440
0.545867025852200
0.367664754390714
0.402774780988691
0.476839542388913
```

xY =

```
xyz: [3x1 double]
name: 'VOI'
Ic: 0
def: 'sphere'
spec: 0
str: '0.0mm sphere'
XYZmm: [3x1 double]
X0: [12x0 double]
y: [12x1 double]
u: [12x1 double]
v: 1
s: 2.400461236262975
```

- YはPlotで書き換えられてしまうため、Excelなどに保存するか、他の変数として保存

Adjustedでグラフを出した後の Y, y, beta

```
>> y
```

```
y =
```

```
-0.020378047748972  
-0.038993441746208  
-0.014650876517858  
 0.010380302671824  
-0.001701707168986  
-0.065838806675451  
 0.027372810870051  
 0.087014141304140  
 0.101779403907900  
-0.076422863052862  
-0.041312837288353  
 0.032751921444774
```

```
>> Y
```

```
Y =
```

```
-0.015996814394991  
-0.044226486856738  
 0.006586923574408  
 0.006586923574408  
-0.015996814394991  
-0.072456159318486  
 0.046108465020855  
 0.079984071974952  
 0.079984071974952  
-0.072456159318486  
-0.044226486856738  
 0.046108465020855
```

- Yとyの微妙な違いはどこから？

```
>> beta
```

```
beta =
```

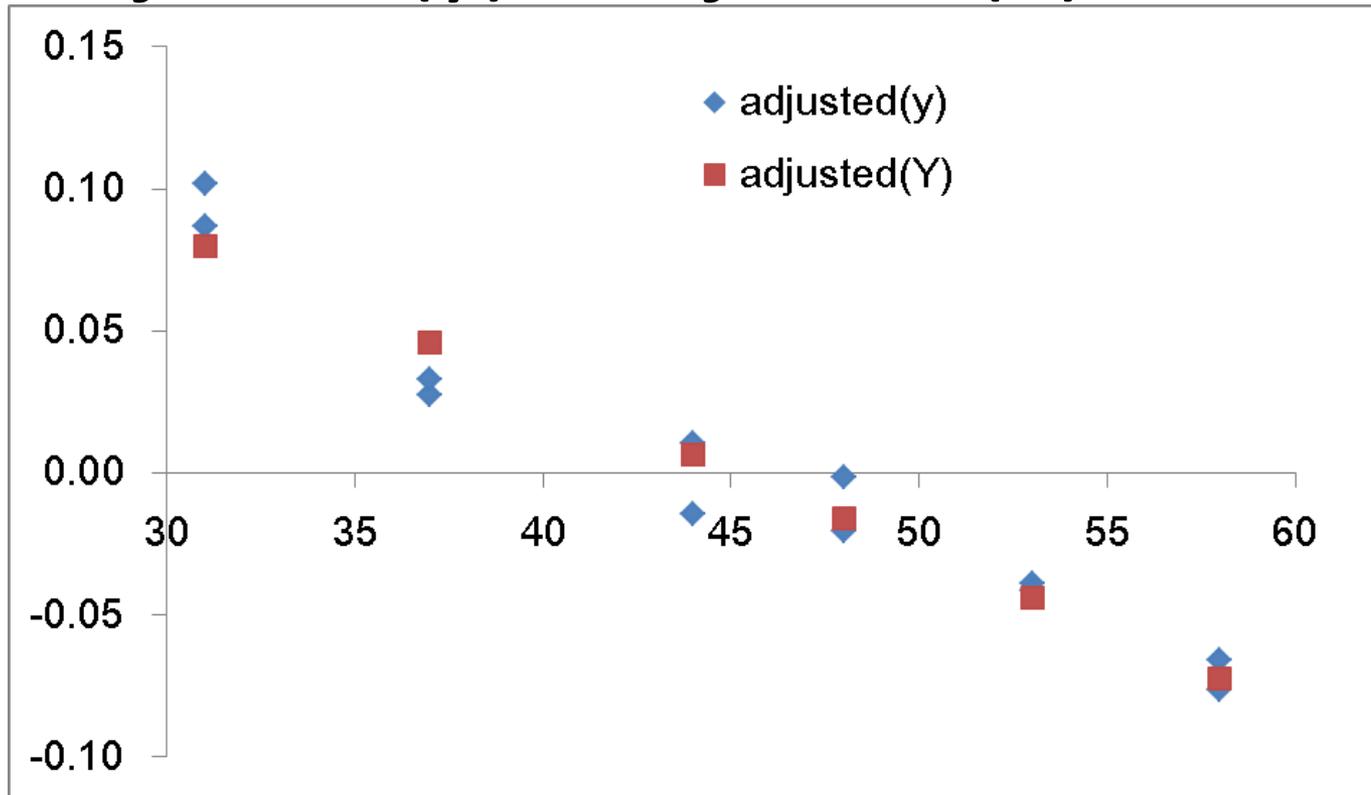
```
 0.6991  
-0.0056
```

VVとadjusted(y), adjusted(Y)の関係

| Age | VV | adjusted(y) | adjusted(Y) | VV- adjusted(y) | VV- adjusted(Y) |
|------|-------|-------------|-------------|--------------------|--------------------|
| 48 | 0.42 | -0.02 | -0.02 | 0.44 | 0.44 |
| 53 | 0.41 | -0.04 | -0.04 | 0.44 | 0.45 |
| 44 | 0.43 | -0.01 | 0.01 | 0.44 | 0.42 |
| 44 | 0.45 | 0.01 | 0.01 | 0.44 | 0.45 |
| 48 | 0.44 | 0.00 | -0.02 | 0.44 | 0.46 |
| 58 | 0.38 | -0.07 | -0.07 | 0.44 | 0.45 |
| 37 | 0.47 | 0.03 | 0.05 | 0.44 | 0.43 |
| 31 | 0.53 | 0.09 | 0.08 | 0.44 | 0.45 |
| 31 | 0.55 | 0.10 | 0.08 | 0.44 | 0.47 |
| 58 | 0.37 | -0.08 | -0.07 | 0.44 | 0.44 |
| 53 | 0.40 | -0.04 | -0.04 | 0.44 | 0.45 |
| 37 | 0.48 | 0.03 | 0.05 | 0.44 | 0.43 |
| mean | 45.17 | 0.44 | 0.00 | 0.44 | 0.44 |

- $\text{adjusted}(y) = VV - VV$ の平均値
- $\text{adjusted}(Y)$ の値は年齢ごとに一致している→推定値

adjusted(y)とadjusted(Y)の違い



- $VV = c + \beta x + \varepsilon$
- $adjusted(y) = VV - mean(VV)$
- $adjusted(Y) = adjusted(y) - \varepsilon = c + \beta x - mean(VV)$

検証：VVとパラメータからadjusted(Y) を算出

| beta | Age | VV | $Y=c+\beta x-\text{mean}$ | adjusted(Y) |
|------------------|-----|------|---------------------------|-------------|
| c: 0.699 | 48 | 0.42 | -0.02 | -0.02 |
| β : -0.006 | 53 | 0.41 | -0.04 | -0.04 |
| | 44 | 0.43 | 0.01 | 0.01 |
| Mean | 44 | 0.45 | 0.01 | 0.01 |
| 0.44 | 48 | 0.44 | -0.02 | -0.02 |
| | 58 | 0.38 | -0.07 | -0.07 |
| | 37 | 0.47 | 0.05 | 0.05 |
| | 31 | 0.53 | 0.08 | 0.08 |
| | 31 | 0.55 | 0.08 | 0.08 |
| | 58 | 0.37 | -0.07 | -0.07 |
| | 53 | 0.40 | -0.04 | -0.04 |
| | 37 | 0.48 | 0.05 | 0.05 |

Predictedでグラフを出した後の Y, y, beta

```
>> y
```

```
y =
```

```
-0.275386088986763  
-0.294001482984000  
-0.269658917755649  
-0.244627738565967  
-0.256709748406777  
-0.320846847913243  
-0.227635230367740  
-0.167993899933651  
-0.153228637329891  
-0.331430904290654  
-0.296320878526145  
-0.222256119793017
```

```
Y =
```

```
-0.271004855632782  
-0.299234528094530  
-0.248421117663383  
-0.248421117663383  
-0.271004855632782  
-0.327464200556278  
-0.208899576216936  
-0.175023969262838  
-0.175023969262838  
-0.327464200556278  
-0.299234528094530  
-0.208899576216936
```

- この微妙な違いはどこから？
- 変数betaはadjusted, predictedともに変わらず

```
>> beta
```

```
beta =
```

```
0.6991  
-0.0056
```

VVとpredicted(y), predicted(Y)の関係

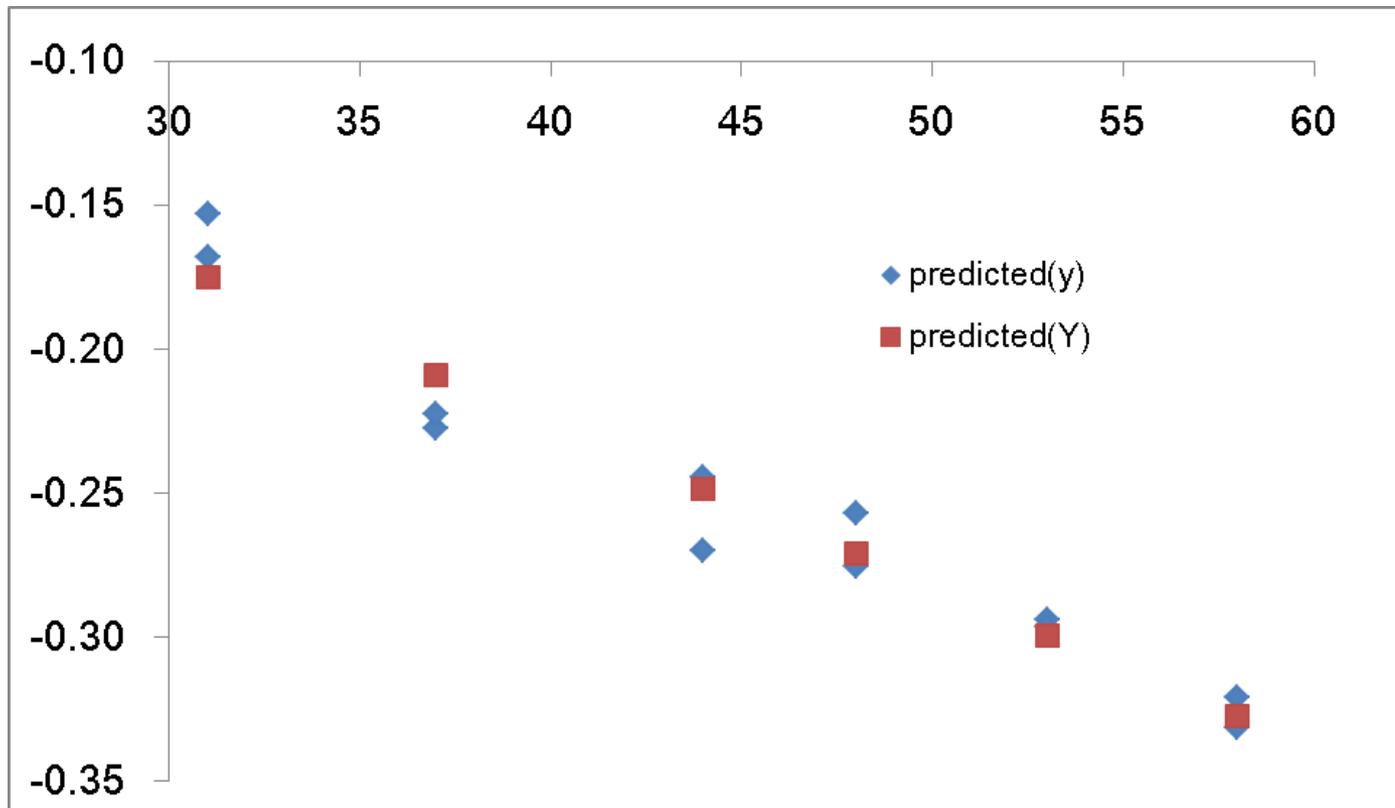
| | Age | VV | predicted(y) | predicted(Y) | VV - predicted(y) | VV - predicted(Y) |
|------|-------|------|--------------|--------------|----------------------|----------------------|
| | 48 | 0.42 | -0.28 | -0.27 | 0.699 | 0.695 |
| | 53 | 0.41 | -0.29 | -0.30 | 0.699 | 0.704 |
| | 44 | 0.43 | -0.27 | -0.25 | 0.699 | 0.678 |
| | 44 | 0.45 | -0.24 | -0.25 | 0.699 | 0.703 |
| | 48 | 0.44 | -0.26 | -0.27 | 0.699 | 0.713 |
| | 58 | 0.38 | -0.32 | -0.33 | 0.699 | 0.706 |
| | 37 | 0.47 | -0.23 | -0.21 | 0.699 | 0.680 |
| | 31 | 0.53 | -0.17 | -0.18 | 0.699 | 0.706 |
| | 31 | 0.55 | -0.15 | -0.18 | 0.699 | 0.721 |
| | 58 | 0.37 | -0.33 | -0.33 | 0.699 | 0.695 |
| | 53 | 0.40 | -0.30 | -0.30 | 0.699 | 0.702 |
| | 37 | 0.48 | -0.22 | -0.21 | 0.699 | 0.686 |
| mean | 45.17 | 0.44 | -0.26 | -0.26 | 0.699 | 0.699 |

- $\text{predicted}(y) = VV - c$
- $\text{predicted}(Y)$ の値はadjusted(Y)と同様に年齢ごとに一致している

beta =

0.6991
-0.0058

predicted(y)とpredicted(Y)の違い



- $VV = c + \beta x + \varepsilon$
- $predicted(y) = VV - c = \beta x + \varepsilon$
- $predicted(Y) = predicted(y) - \varepsilon = \beta x$

検証：VVとパラメータから predicted(y)とpredicted(Y)を算出

| beta | Age | raw | $y=VV-c$ | predicted(y) | $Y=\beta x$ | predicted(Y) |
|-------------|-----|------|----------|--------------|-------------|--------------|
| 0.699 | 48 | 0.42 | -0.28 | -0.28 | -0.27 | -0.27 |
| -0.006 | 53 | 0.41 | -0.29 | -0.29 | -0.30 | -0.30 |
| | 44 | 0.43 | -0.27 | -0.27 | -0.25 | -0.25 |
| mean | 44 | 0.45 | -0.24 | -0.24 | -0.25 | -0.25 |
| 0.44 | 48 | 0.44 | -0.26 | -0.26 | -0.27 | -0.27 |
| | 58 | 0.38 | -0.32 | -0.32 | -0.33 | -0.33 |
| | 37 | 0.47 | -0.23 | -0.23 | -0.21 | -0.21 |
| | 31 | 0.53 | -0.17 | -0.17 | -0.18 | -0.18 |
| | 31 | 0.55 | -0.15 | -0.15 | -0.18 | -0.18 |
| | 58 | 0.37 | -0.33 | -0.33 | -0.33 | -0.33 |
| | 53 | 0.40 | -0.30 | -0.30 | -0.30 | -0.30 |
| | 37 | 0.48 | -0.22 | -0.22 | -0.21 | -0.21 |

Plotのまとめ

- Plotで得られる値は、adjustedとpredictedがあり、それぞれに y (adjusted)と Y (fitted)がある。
 Y (fitted)は y (adjusted)から誤差 ε をひいたもの。
つまり回帰直線においては回帰直線上の推定値。
- 回帰直線においては、
 - Adjusted: ボクセル値(VV)からボクセル値の平均(mean)をひいたものである
 - $\text{Adjusted}(y) = VV - \text{mean}$
 - $\text{Adjusted}(Y) = VV - \text{mean} - \varepsilon$
 - Predicted: ボクセル値(VV)から定数 c をひいたものである
 - $\text{Predicted}(y) = VV - c$
 - $\text{Predicted}(Y) = VV - c - \varepsilon = \beta x$