## 【統計一口メモ　第 39 話】

## ＜最小二乗平均とは？＞

名古屋市立大学大学院医学研究科 非常勤講師　薬学博士　松本一彦

ある治療薬の添付文書に「最小二乗平均(least square means，LSM)とその 95％信頼区間」が載っていました。基礎研究の毒性家には、ほとんど目にすることのない用語です。でも、一般人も目にする添付文書に載っているこの用語の意味を知っておかねばーーです。

表1　12週時の1日平均排尿回数の変化量

| 投与群 | 症例数 | 投与前 | 変化量# | プラセボ群との比較## |
|---|---|---|---|---|
| プラセボ群 | 369 | 11.20±2.40 | −1.21(−1.40，−1.03) | − |
| A剤50mg群 | 370 | 11.13±2.37 | −2.08(−2.27，−1.89) | p＜0.0001 |
| A剤100mg群 | 368 | 11.08±2.25 | −2.03(−2.22，−1.84) | p＜0.0001 |

平均値±標準偏差
＃最小二乗平均値(95％信頼区間)
##有意水準両側5％

### §1. 算術平均と最小二乗平均の違い

“通常の平均値”は線形モデル y＝aX＋b の a が 0 のとき、すなわち傾きがないときに最小二乗法を用いてパラメータであるbで計算します。一方、“最小二乗平均”の場合は経時変化のように、投与前を X としたとき、すなわち a≠0 のときに用います。したがって、表1のようなプラセボ群と2つの投与群で変化量の違いを求める場合、線形モデルとして 3 本の平行な回帰直線の間隔を比較することになります。すなわち、経時変化のあるデータでは傾きのない 3 群の平均値の比較と比べて、傾きを考慮して比較をする方が正しい解析といえます。

それでは、2 群比較の例題を使って「最小二乗平均とその 95％信頼区間の求め方」を見ていきましょう。 ただ、計算には通常みかけない行列式がふんだんに使われています。実際には、エクセル関数(MMULT、TRANSPOSE、MINVERSE)で処理できるので、行列式の詳しいことを知らなくても計算することができるのですが、じっくり取り組んでみたい方は、高橋行雄「最尤法によるポアソン回帰分析入門」[1)]を参照してください。

### §2. 最小二乗平均を求める

例題：血圧降下剤AとBの拡張期血圧を投与前と投与1週後の変化量を調べた。両群の最小二乗平均と95％信頼区間を求め、差を検定したい。

例題として2群比較をとりあげます。最小二乗平均の求め方と従来行われている前後値の 2 群比較との違いを見てみましょう。

表2. 薬剤 A と薬剤Bの投与前後の血圧値

| 薬剤A | | 薬剤B | |
|---|---|---|---|
| 前値 | 後値 | 前値 | 後値 |
| 148 | 158 | 184 | 197 |
| 144 | 152 | 166 | 157 |
| 146 | 160 | 139 | 127 |
| 163 | 174 | 158 | 150 |
| 159 | 160 | 148 | 135 |
| 157 | 167 | 174 | 176 |
| 155 | 151 | 151 | 145 |
| 165 | 161 | 167 | 180 |
| 177 | 169 | 165 | 155 |
| 156 | 160 | 154 | 148 |

## 表3. 最小二乗平均計算原表

| | デザイン行列 | | | | | |
|---|---|---|---|---|---|---|
| | | | 前値 | 後値 | 差 | 推定値 |
| | $x_0$ | $a$ | $X$ | $Y$ | $D$ | $D$^ |
| | 1 | 0 | 148 | 158 | 10 | 2.8 |
| | 1 | 0 | 144 | 152 | 8 | 2.1 |
| | 1 | 0 | 146 | 160 | 14 | 2.4 |
| | 1 | 0 | 163 | 174 | 11 | 5.2 |
| 薬剤A | 1 | 0 | 159 | 160 | 1 | 4.5 |
| | 1 | 0 | 157 | 167 | 10 | 4.2 |
| | 1 | 0 | 155 | 151 | -4 | 3.9 |
| | 1 | 0 | 165 | 161 | -4 | 5.5 |
| | 1 | 0 | 177 | 169 | -8 | 7.4 |
| | 1 | 0 | 156 | 160 | 4 | 4.0 |
| | 1 | 1 | 184 | 197 | 13 | 0.1 |
| | 1 | 1 | 166 | 157 | -9 | -2.7 |
| | 1 | 1 | 139 | 127 | -12 | -7.1 |
| | 1 | 1 | 158 | 150 | -8 | -4.0 |
| | 1 | 1 | 148 | 135 | -13 | -5.6 |
| 薬剤B | 1 | 1 | 174 | 176 | 2 | -1.5 |
| | 1 | 1 | 151 | 145 | -6 | -5.1 |
| | 1 | 1 | 167 | 180 | 13 | -2.6 |
| | 1 | 1 | 165 | 155 | -10 | -2.9 |
| | 1 | 1 | 154 | 148 | -6 | -4.7 |
| | 1 | 0 | 158.80 | ←2群前値平均 | | 4.5 |
| 薬剤A | 平均 | $\mu$^$_1$ | 157.0 | 161.2 | 4.2 | 最小2乗平均 |
| | *SD* | $\sigma$^$_1$ | 9.9 | 7.2 | 7.6 | |
| | 1 | 1 | 158.80 | ←2群前値平均 | | -3.89 |
| 薬剤B | 平均 | $\mu$^$_2$ | 160.6 | 157.0 | -3.6 | 最小2乗平均 |
| | *SD* | $\sigma$^$_2$ | 13.3 | 21.5 | 9.7 | |
| 平均の差 | $\mu$^$_2$ - $\mu$^$_1$ | | 3.60 | -4.20 | -7.80 | -8.38 |
| | 差の*SE* | | 5.23 | 7.16 | 3.89 | 3.96 |
| | $t$値 | | 0.688 | -0.587 | -2.007 | -2.117 |
| | 両側$p$値 | | 0.500 | 0.565 | 0.060 | 0.0493 |

| Dbar | Di-Dbar | D^ | D^-Dbar | D-D^ |
|---|---|---|---|---|
| 4.2 | 5.8 | 2.8 | -1.4 | 7.2 |
| 4.2 | 3.8 | 2.1 | -2.1 | 5.9 |
| 4.2 | 9.8 | 2.4 | -1.8 | 11.6 |
| 4.2 | 6.8 | 5.2 | 1.0 | 5.8 |
| 4.2 | -3.2 | 4.5 | 0.3 | -3.5 |
| 4.2 | 5.8 | 4.2 | 0.0 | 5.8 |
| 4.2 | -8.2 | 3.9 | -0.3 | -7.9 |
| 4.2 | -8.2 | 5.5 | 1.3 | -9.5 |
| 4.2 | -12.2 | 7.4 | 3.2 | -15.4 |
| 4.2 | -0.2 | 4.0 | -0.2 | 0.0 |
| -3.6 | 16.6 | 0.1 | 3.7 | 12.9 |
| -3.6 | -5.4 | -2.7 | 0.9 | -6.3 |
| -3.6 | -8.4 | -7.1 | -3.5 | -4.9 |
| -3.6 | -4.4 | -4.0 | -0.4 | -4.0 |
| -3.6 | -9.4 | -5.6 | -2.0 | -7.4 |
| -3.6 | 5.6 | -1.5 | 2.1 | 3.5 |
| -3.6 | -2.4 | -5.1 | -1.5 | -0.9 |
| -3.6 | 16.6 | -2.6 | 1.0 | 15.6 |
| -3.6 | -6.4 | -2.9 | 0.7 | -7.1 |
| -3.6 | -2.4 | -4.7 | -1.1 | -1.3 |
| | 1360.00 | | 63.18 | 1296.82 |
| | ST | | SR | Se |
| 自由度 | 19 | | 2 | 17 |

表 3 はこれから求める最小二乗平均の計算に必要な項目とその結果です。ここでは、両群のデータから**"後値Y"**の 2 群平均値の差では有意差はみられず(P=0.565)**"前後差D"**では p 値が p=0.060 と棄却限界値(0.05)に近い値を示している例をとりあげています。前値を共変量とする最小二乗平均を用いた 2 群の差の検定では p=0.0493 となっていますが、どのようにして求めたのかを見てみましょう。

最小二乗平均は、**デザイン行列と係数(β^)の行列**から求めます。それぞれの求め方をひも解いていきましょう。 デザイン行列はここでは「最小二乗平均算出のための**行列 X**」を意味します。

表4

| | | AK | AL | AM | AN |
|---|---|---|---|---|---|
| | 最小二乗平均算出のための行列X | | | | |
| | | $X_0$ | a | X値2群総平均 | 係数(β^) |
| 9 | 薬剤A | 1 | 0 | 158.8 | -20.939 |
| 10 | 薬剤B | 1 | 1 | 158.8 | -8.376 |
| 11 | | | | | 0.160 |

MMULT(X の範囲、係数の範囲=MMULT(AK9:AM9,AN9:AN11)

このような行列関数 MMULT による計算から表5のような、最小二乗値を求めることができますが、

表 5

| | 最小二乗平均 |
|---|---|
| 薬剤A | 4.49 |
| 薬剤B | -3.89 |

問題は“行列 X”と“係数（$\beta$＾）”行列の求め方です。
それでは、ひとつずつ手順を追ってみましょう。

手順1. “投与前値の2群総平均”とした 2 群の行列 X を作成する

| 群名 | $x_0$ | a | X値2群総平均 |
|---|---|---|---|
| 薬剤A＝ | 1 | 0 | 158.8 |
| 薬剤B＝ | 1 | 1 | 158.8 |

$x_0$ に1、a に1と0を配置します。これをダミー変数と呼びます。
今回は薬剤 A を1，0　薬剤Bを1，1としました。その横に「前値 X の 2 群をまとめた総平均値」を配置します。これを“**行列 X**”と呼びます。

手順2. 係数（$\beta$＾）行列を作成する

| 係数（$\beta$＾）= $(X^TX)^{-1}X^TD$ | | |
|---|---|---|
| -20.939 | | |
| -8.376 | | |
| 0.16 | | |

①「行列 X の転置＆逆行列 $(X^TX)^{-1}$」と
②「行列 X と前後差 D 値の転置行列 $X^TD$」から求めます。

MMULT（$(X^TX)^{-1}$ の範囲、$X^TD$ の範囲）
行と列を入れ替えた行列を**転置行列**といい$X^TX$と表します。エクセル関数は TRANSPOSE です。
行列式を割り算して求める行列を**逆行列**といい$(X^TX)^{-1}$と表します。エクセル関数は MINVERSE です。
①行列 X の転置＆逆行列$(X^TX)^{-1}$ 作成
MINVERSE(MMULT(TRANSPOSE(X の範囲)、X の範囲)
＜Xの範囲＞
最小二乗平均算出のための行列 X

| $x_0$ | $a$ | $X_{ij}$ |
|---|---|---|
| 1 | 0 | 148 |
| 1 | 0 | 144 |
| 1 | 0 | 146 |
| 1 | 0 | 163 |
| 1 | 0 | 159 |
| 1 | 0 | 157 |
| 1 | 0 | 155 |
| 1 | 0 | 165 |
| 1 | 0 | 177 |
| 1 | 0 | 156 |
| 1 | 1 | 184 |
| 1 | 1 | 166 |
| 1 | 1 | 139 |
| 1 | 1 | 158 |
| 1 | 1 | 148 |
| 1 | 1 | 174 |
| 1 | 1 | 151 |
| 1 | 1 | 167 |
| 1 | 1 | 165 |
| 1 | 1 | 154 |

$(X^TX)^{-1}$

| | | | |
|---|---|---|---|
| $\hat{\beta}_0$ | 10.1020 | 0.1293 | -0.0637 |
| $\hat{\beta}_1$ | 0.1293 | 0.2053 | -0.0015 |
| $\hat{\beta}_2$ | -0.0637 | -0.0015 | 0.0004 |

②行列 X と前後差 D 値の転置行列 $X^TD$ 作成

MMULT(TRANSPOSE(X の範囲)、D の範囲)

＜X の範囲＞　　　＜Dの範囲＞

最小二乗平均算出のための行列 X

| $x_0$ | a | $X_{ij}$ |
|---|---|---|
| 1 | 0 | 148 |
| 1 | 0 | 144 |
| 1 | 0 | 146 |
| 1 | 0 | 163 |
| 1 | 0 | 159 |
| 1 | 0 | 157 |
| 1 | 0 | 155 |
| 1 | 0 | 165 |
| 1 | 0 | 177 |
| 1 | 0 | 156 |
| 1 | 1 | 184 |
| 1 | 1 | 166 |
| 1 | 1 | 139 |
| 1 | 1 | 158 |
| 1 | 1 | 148 |
| 1 | 1 | 174 |
| 1 | 1 | 151 |
| 1 | 1 | 167 |
| 1 | 1 | 165 |
| 1 | 1 | 154 |

| 前後差 |
|---|
| $D_{ij}$ |
| 10 |
| 8 |
| 14 |
| 11 |
| 1 |
| 10 |
| -4 |
| -4 |
| -8 |
| 4 |
| 13 |
| -9 |
| -12 |
| -8 |
| -13 |
| 2 |
| -6 |
| 13 |
| -10 |
| -6 |

$X^TD$

| | |
|---|---|
| $\hat{\beta}_0$ | 6 |
| $\hat{\beta}_1$ | -36 |
| $\hat{\beta}_2$ | 1207 |

③係数（β＾）を求める

MMULT（$(X^TX)^{-1}$ の範囲、$X^TD$ の範囲）

①と②で求めた値から算出します。

| | $(X^TX)^{-1}$の範囲 | | | $X^TD$の範囲 |
|---|---|---|---|---|
| $\hat{\beta}_0$ | 10.1020 | 0.1293 | -0.0637 | 6 |
| $\hat{\beta}_1$ | 0.1293 | 0.2053 | -0.0015 | -36 |
| $\hat{\beta}_2$ | -0.0637 | -0.0015 | 0.0004 | 1207 |

係数（β＾）＝$(X^TX)^{-1}X^TD$

| |
|---|
| −20.939 |
| −8.376 |
| 0.160 |

手順3. 最小二乗平均を求める

表 4、表 5を再掲します。

MMULT(X の範囲、係数の範囲)=MMULT(AK9:AM9,AN9:AN11)

＜X の範囲＞　　　＜係数の範囲＞

最小二乗平均算出のための行列 X

| | | AK | AL | AM | AN |
|---|---|---|---|---|---|
| | | x | a | X値2群総平均 | 係数（β＾） |
| 9 | 薬剤A | 1 | 0 | 158.8 | -20.9390 |
| 10 | 薬剤B | 1 | 1 | 158.8 | -8.3760 |
| 11 | | | | | 0.1600 |

手順1. に順ずる

| | |
|---|---|
| | 最小二乗平均 |
| 薬剤A | 4.49 |
| 薬剤B | -3.89 |

## §3. 最小二乗平均の95%信頼区間を求める

95%信頼区間を算出するためには、標準誤差(SE)を求めなければなりません。
SEは√分散Var(D)なので、まず手順1で分散を求めます。

手順1. 前後差D値に対する**分散Var(D)**を求める。

$Var(D) = L\Sigma(\hat{\beta})L^T$

$=$[行列X]x[共分散行列$\Sigma(\hat{\beta})$]x[デザイン転置行列$X^T$]

薬剤A　<行列X>　<共分散行列>

| x | a | X値の2群総平均 | 共分散行列$\Sigma(\hat{\beta})=(X^TX)^{-1}\hat{\sigma^2}$ | | |
|---|---|---|---|---|---|
| 1 | 0 | 158.8 | 770.6164 | 9.8669 | -4.8598 |
| | | | 9.8669 | 15.6578 | -0.1114 |
| | | | -4.8598 | -0.1114 | 0.0310 |

§2 手順1に順ずる

①共分散行列を求める。

ここで、新しく**共分散行列**$\Sigma(\hat{\beta})$が出てきました。

$\Sigma(\hat{\beta})=(X^TX)^{-1}\hat{\sigma^2}$

さらに、式には$\hat{\sigma^2}$も登場しています。

②**誤差分散**の推定値$\hat{\sigma^2}$を求める

$\hat{\sigma^2}$は誤差分散の推定値といって、**残差平方和**Se/**残差自由度**から求めます。

残差自由度は表3のSe自由度から求めます。今回は17となります。

③残差平方和Seを求める。

$Se = \hat{\varepsilon}^T\hat{\varepsilon}$

MMULT(TRANSPOSE(ε^の範囲)、ε^の範囲)

ε^は**残差ベクトル**といいます。

④残差ベクトルε^を求める。

$\hat{\varepsilon} = D - \hat{D}$ ＝(Dの範囲－D^の範囲)

D推定値**D^**はMMULT(行列Xの範囲、係数β^の範囲)で求める。

最小二乗平均算出のための行列 X

| $x_0$ | $a$ | $X_{ij}$ |
|---|---|---|
| 1 | 0 | 148 |
| 1 | 0 | 144 |
| 1 | 0 | 146 |
| 1 | 0 | 163 |
| 1 | 0 | 159 |
| 1 | 0 | 157 |
| 1 | 0 | 155 |
| 1 | 0 | 165 |
| 1 | 0 | 177 |
| 1 | 0 | 156 |
| 1 | 1 | 184 |
| 1 | 1 | 166 |
| 1 | 1 | 139 |
| 1 | 1 | 158 |
| 1 | 1 | 148 |
| 1 | 1 | 174 |
| 1 | 1 | 151 |
| 1 | 1 | 167 |
| 1 | 1 | 165 |
| 1 | 1 | 154 |

| 係数($\beta$^) |
|---|
| -20.9389 |
| -8.3764 |
| 0.1601 |

| $D$^ |
|---|
| 2.8 |
| 2.1 |
| 2.4 |
| 5.2 |
| 4.5 |
| 4.2 |
| 3.9 |
| 5.5 |
| 7.4 |
| 4.0 |
| 0.1 |
| -2.7 |
| -7.1 |
| -4.0 |
| -5.6 |
| -1.5 |
| -5.1 |
| -2.6 |
| -2.9 |
| -4.7 |

| $D_{ij}$ | $D$^ |
|---|---|
| 10 | 2.8 |
| 8 | 2.1 |
| 14 | 2.4 |
| 11 | 5.2 |
| 1 | 4.5 |
| 10 | 4.2 |
| -4 | 3.9 |
| -4 | 5.5 |
| -8 | 7.4 |
| 4 | 4.0 |
| 13 | 0.1 |
| -9 | -2.7 |
| -12 | -7.1 |
| -8 | -4.0 |
| -13 | -5.6 |
| 2 | -1.5 |
| -6 | -5.1 |
| 13 | -2.6 |
| -10 | -2.9 |
| -6 | -4.7 |

| $\varepsilon$^ |
|---|
| $D$-$D$^ |
| 7.2 |
| 5.9 |
| 11.6 |
| 5.8 |
| -3.5 |
| 5.8 |
| -7.9 |
| -9.5 |
| -15.4 |
| 0.0 |
| 12.9 |
| -6.3 |
| -4.9 |
| -4.0 |
| -7.4 |
| 3.5 |
| -0.9 |
| 15.6 |
| -7.1 |
| -1.3 |

＜誤差平方和＞

残差平方和 Se= $\varepsilon\hat{}^{\,T}\varepsilon\hat{}$

MMULT(TRANSPOSE(ε^の範囲)、ε^の範囲)

| 1296.8 |
|---|

＜誤差分散＞

誤差分散($\sigma^{2}$^)＝Se/残差自由度

＝1296.8/17＝76.3

| 76.3 |
|---|

ようやく「共分散行列Σ($\beta$^)=$(X^TX)^{-1}\sigma^{2\wedge}$」が仕上がりました。

| 共分散行列　Σ($\beta$^)=$(X^TX)^{-1}\sigma^{2\wedge}$ | | |
|---|---|---|
| 770.616 | 9.867 | -4.860 |
| 9.867 | 15.658 | -0.111 |
| -4.860 | -0.111 | 0.031 |

手順2. ここで手順1に戻ります。

　分散 Var(D)を求める。 Var(D)=LΣ($\beta$^)$L^T$

Var(D)＝「行列 X」x「共分散行列 Σ($\beta$^)」x [行列 X の転置行列 $X^T$]

薬剤 A

| x | a | X値の2群総平均 | 共分散行列Σ(β^)=($X^TX$)$^{-1}$σ$^{2}$^ | | |
|---|---|---|---|---|---|
| 1 | 0 | 158.8 | 770.6164 | 9.8669 | -4.8598 |
| | | | 9.8669 | 15.6578 | -0.1114 |
| | | | -4.8598 | -0.1114 | 0.0310 |

手順1に順ずる

薬剤 B

| x | a | X値の2群総平均 | 共分散行列Σ(β^)=($X^TX$)$^{-1}$σ$^{2}$^ | | |
|---|---|---|---|---|---|
| 1 | 1 | 158.8 | 770.6164 | 9.8669 | -4.8598 |
| | | | 9.8669 | 15.6578 | -0.1114 |
| | | | -4.8598 | -0.1114 | 0.0310 |

手順1に順ずる

| | 分散V |
|---|---|
| 薬剤A＝ | 7.7286 |
| 薬剤B＝ | 7.7286 |

=MMULT(MMULT(行列 X の範囲、共分散行列の範囲),TRANSPOSE(行列 X の範囲))＝7.7286

手順3. SE＝(√分散)を求める

| | SE |
|---|---|
| 薬剤A＝ | 2.7800 |
| 薬剤B＝ | 2.7800 |

‘=SQRT(分散)
‘=SQRT(分散)

手順4. t0.05(残差自由度) x SE を求める。$t_{0.05}$(17)=2.10

| | $t_{0.05}$ x SE |
|---|---|
| 薬剤A＝ | 5.8654 |
| 薬剤B＝ | 5.8654 |

=TINV(0.05,残差自由度)*SE
=TINV(0.05,残差自由度)*SE

手順5. 95%信頼区間＝最小二乗平均±t0.05 x SE

| | L95% | U95% |
|---|---|---|
| 薬剤A＝ | -1.40 | 10.33 |
| 薬剤B＝ | -9.77 | 1.96 |

まとめ　　前後差 D の最小二乗平均値と 95%信頼区間表

| | | 95%信頼区間 | |
|---|---|---|---|
| 群名 | 最小二乗平均 | 下側95% | 上側95% |
| 薬剤A＝ | 4.49 | -1.40 | 10.33 |
| 薬剤B＝ | -3.89 | -9.77 | 1.96 |

## §4. 薬剤 A と薬剤 B の最小二乗平均の差の検定

手順1. 両群の差の分散の平方根 SE は回帰分析のダミー変数 a の推定値 $\beta_1$^であることから
共分散行列の分散 15.658 から√して標準誤差 SE＝3.957 を得る。

| | 共分散行列 $\Sigma(\beta\hat{})=(X^TX)^{-1}\sigma^{2}\hat{}$ | | |
|---|---|---|---|
| $\beta_0\hat{}$ | 770.616 | 9.867 | -4.860 |
| $\beta_1\hat{}$ | 9.867 | $\boxed{15.658}$ | -0.111 |
| $\beta_2\hat{}$ | -4.860 | -0.111 | 0.031 |

手順2. 最小二乗平均の差＝4.49-(-3.89)＝8.376

手順3. t値＝最小二乗平均の差/SE＝8.376/3.957＝2.117

手順4. p値＝p値＝T.DIST.2T(t値、残差自由度)＝T.DIST.2T(2.117,17)＝0.0493

| | | | | 95%信頼区間 | | |
|---|---|---|---|---|---|---|
| | 最小二乗平均 | 標準誤差 | t値 | 下限 | 上限 | p値 |
| 薬剤A,B差 | 8.376 | 3.957 | 2.117 | 0.027 | 16.725 | 0.049 |

結果：薬剤 A と薬剤 B の最小二乗平均の差は p=0.0493 で有意差がみられる。

以上のように、行列計算を駆使してようやく、添付文書に書かれていた変化量の最小二乗平均とその 95%信頼区間および群間のp値を求めることができました。最小二乗平均がほとんどの教科書に取り上げられていないことも、その解析手順の複雑さによっているのかもしれません。なお、実践で最小二乗平均を使うにはJMPソフトを使用することをお薦めします。

注） 行列式は、厳密には、determinant ｜正方行列｜ を意味します。Excel では、=MDETERM() 関数を使います。行列を使った計算は、matrix algebra 行列代数 と言いますが、ここでは、行列を行列計算 または行列計算式 とすることを推奨します（高橋行雄コメント）。

謝辞：行列計算については、高橋行雄氏にご指導を受けました。また、本メモについても査読をお願いして大事な注も入れていただきました。御礼申し上げます。

以上

1) 高橋行雄「最尤法によるポアソン回帰分析入門」 カクワークス社
2) Pharmaco Basic Ver.17　Pharmaco 工房（https://pharmaco club/）