線型推測論

第08回 信頼区間・一般化最小二乗推定量

千葉大学大学院医学研究院
長島 健悟

信頼区間

- $\mathbf{c}'\boldsymbol{\beta}$の信頼区間は比較的用いられる
  - 理論的には単一仮説の検定統計量をそのまま流用できる
- $\dfrac{\mathbf{c}'\widehat{\boldsymbol{\beta}}}{\sqrt{s^2\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim t(n-p)$

$$\mathbf{c}'\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2,n-p}\sqrt{s^2\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}$$

誤差の条件の一般化

- 今まで我々は誤差の三条件のもとで
  - 最小二乗推定量
  - 分散推定量
  - 仮説検定 (正規性の仮定を追加)
- 今度は誤差の三条件のうちの一部の制約を緩めた場合を議論する

誤差の条件の一般化

- 新しい誤差の条件
  - $\mathrm{Cov}[\boldsymbol{\epsilon}] = \sigma^2\mathbf{V}$：$\mathbf{V}$は既知の正定値対称行列とする
  - $\mathrm{Cov}[\boldsymbol{\epsilon}] \neq \sigma^2\mathbf{I}$の場合を許し, 誤差の間に相関がある場合の推測を考える

- 系8-1
  - $\mathrm{Cov}[\mathbf{Y}] = \sigma^2\mathbf{V}$

新しい条件の整理

- 新しい条件の下では以下が成り立つ
  - $\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\epsilon}$
  - $\mathrm{E}[\boldsymbol{\epsilon}] = \mathbf{0}$
  - $\mathrm{E}[\mathbf{Y}] = \mathbf{X\beta}$
  - $\mathrm{Cov}[\boldsymbol{\epsilon}] = \mathrm{Cov}[\mathbf{Y}] = \sigma^2\mathbf{V}$

一般化最小二乗推定量

- 定理8-2
  - $\boldsymbol{\beta}$のBLUEは

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

  である

一般化最小二乗推定量

- 定理8-3
  - $\widehat{\boldsymbol{\beta}}$の分散共分散行列は

$$\mathrm{Cov}\left(\widehat{\boldsymbol{\beta}}\right) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$$

  である

一般化最小二乗推定量

- 定理8-4
  - σ^2の不偏推定量は

$$s^2 = \frac{\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right)}{n-p}$$

$$= \frac{\mathbf{Y}'[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{Y}}{n-p}$$

である

演習問題：定理8-2

- 証明のアウトライン
  - $\mathbf{V}$は既知の正定値対称行列である
  - $\mathbf{V} = \mathbf{P}\mathbf{P}'$にコレスキー分解できるうえに$\mathbf{P}$は正則行列になる
  - $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$に左から$\mathbf{P}^{-1}$をかけると・・・

$$\mathbf{P}^{-1}\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{P}^{-1}\boldsymbol{\epsilon}$$

演習問題：定理8-2

- $\mathrm{E}[\mathbf{P}^{-1}\boldsymbol{\epsilon}] = ?$

$$\mathrm{E}[\mathbf{P}^{-1}\boldsymbol{\epsilon}] = \mathbf{P}^{-1}\mathrm{E}[\boldsymbol{\epsilon}] = \mathbf{0}$$

演習問題：定理8-2

- $\mathrm{Cov}[\mathbf{P}^{-1}\boldsymbol{\epsilon}] = ?$

$$
\begin{aligned}
\mathrm{Cov}[\mathbf{P}^{-1}\boldsymbol{\epsilon}] = \mathbf{P}^{-1}\mathrm{Cov}[\boldsymbol{\epsilon}](\mathbf{P}^{-1})' \\
= \mathbf{P}^{-1}\sigma^2\mathbf{V}(\mathbf{P}^{-1})' \\
= \sigma^2\mathbf{P}^{-1}\mathbf{V}(\mathbf{P}^{-1})' \\
= \sigma^2\mathbf{P}^{-1}\mathbf{P}\mathbf{P}'(\mathbf{P}')^{-1} \\
= \sigma^2\mathbf{I}
\end{aligned}
$$

演習問題：定理8-2

- $\mathbf{P}^{-1}\boldsymbol{\epsilon}$は期待値が$\mathbf{0}$で分散が$\sigma^2\mathbf{I}$
  - 誤差の三条件を満たす
- したがって, 式$\mathbf{P}^{-1}\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{P}^{-1}\boldsymbol{\epsilon}$に基づく最小二乗推定量はBLUEを与える

演習問題：定理8-2

- $\widehat{\boldsymbol{\beta}} = ?$

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= [(\mathbf{P}^{-1}\mathbf{X})'(\mathbf{P}^{-1}\mathbf{X})]^{-1}(\mathbf{P}^{-1}\mathbf{X})'(\mathbf{P}^{-1}\mathbf{Y})\\ &= [\mathbf{X}'(\mathbf{P}^{-1})'(\mathbf{P}^{-1})\mathbf{X}]^{-1}\mathbf{X}'(\mathbf{P}^{-1})'(\mathbf{P}^{-1})\mathbf{Y}\\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}\end{aligned}$$

- $$\begin{aligned}(\mathbf{P}^{-1})'(\mathbf{P}^{-1}) &= (\mathbf{P}')^{-1}\mathbf{P}^{-1}\\ &= (\mathbf{P}\mathbf{P}')^{-1}\\ &= \mathbf{V}^{-1}\end{aligned}$$

演習問題：定理8-2

- 前ページのヒント
- 正則行列の積の逆行列
  - $\mathbf{A}, \mathbf{B}$が正則行列ならば
  - $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- 逆行列と転置の交換
  - $\mathbf{A}$が正則行列ならば
  - $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$

演習問題：定理8-3

- $\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = ?$

$$\begin{aligned}\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) &= \mathrm{Cov}[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}] \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathrm{Cov}(\mathbf{Y}) \\ &\qquad [(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]' \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\sigma^2\mathbf{V} \\ &\qquad \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\end{aligned}$$

演習問題：定理8-4

- $\mathbf{P}^{-1}\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{P}^{-1}\boldsymbol{\epsilon}$に基づいて$\sigma^2$の不偏推定量を求める

$$s^2 = \frac{\left(\mathbf{P}^{-1}\mathbf{Y} - \mathbf{P}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{P}^{-1}\mathbf{Y} - \mathbf{P}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{n-p}$$

$$= \frac{\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'(\mathbf{P}^{-1})'\mathbf{P}^{-1}\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)}{n-p}$$

$$= \frac{\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)}{n-p}$$

演習問題：定理8-4

- $\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)'\mathbf{V}^{-1}\left(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}\right)=$
  $\mathbf{Y}'[\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{Y}$
- $\mathbf{A}=\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$とおいておく

演習問題：定理8-4

- $\mathrm{E}[\mathbf{Y}'\mathbf{A}\mathbf{Y}] = ?$

$$
\begin{aligned}
&\mathrm{E}[\mathbf{Y}'\mathbf{A}\mathbf{Y}] \\
&= \mathrm{tr}(\mathbf{A}\mathrm{E}[\mathbf{Y}\mathbf{Y}']) \\
&= \mathrm{tr}(\mathbf{A}\sigma^2\mathbf{V} + \mathbf{A}\mathbf{E}[\mathbf{Y}]\mathbf{E}[\mathbf{Y}]') \\
&= \sigma^2\mathrm{tr}(\mathbf{A}\mathbf{V}) + \mathrm{tr}(\mathbf{E}[\mathbf{Y}]'\mathbf{A}\mathbf{E}[\mathbf{Y}])
\end{aligned}
$$

演習問題：定理8-4

- $\sigma^2 \mathrm{tr}(\mathbf{AV}) = ?$

$$\sigma^2 \mathrm{tr}(\mathbf{AV}) = \sigma^2 \mathrm{tr}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]$$
$$= \sigma^2 \mathrm{tr}(\mathbf{I}_n) - \sigma^2 \mathrm{tr}[\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]$$
$$= \sigma^2 \mathrm{tr}(\mathbf{I}_n) - \sigma^2 \mathrm{tr}[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}]$$
$$= \sigma^2 (n - p)$$

- $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$：$p \times p$次元

演習問題：定理8-4

- $\mathrm{tr}(\mathbf{E[Y]'AE[Y]}) = ?$

$$\mathrm{tr}(\mathbf{E[Y]'AE[Y]}) = \mathbf{E[Y]'AE[Y]}$$
$$= \boldsymbol{\beta}'\mathbf{X}'[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{X}\boldsymbol{\beta}$$
$$= \boldsymbol{\beta}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}$$
$$= 0$$

  - $\mathrm{tr}(\mathbf{E[Y]'AE[Y]})$はスカラー

演習問題：定理8-4

- $\mathrm{E}[s^2] =?$

$$\frac{\mathrm{E}[\mathbf{Y}'\mathbf{A}\mathbf{Y}]}{n-p} = \frac{\sigma^2 \mathrm{tr}(\mathbf{A}\mathbf{V}) + \mathrm{tr}(\mathbf{E}[\mathbf{Y}]'\mathbf{A}\mathbf{E}[\mathbf{Y}])}{n-p}$$

$$= \frac{\sigma^2(n-p)+0}{n-p}$$

$$= \sigma^2$$

- よって不偏である

一般化最小二乗推定量のまとめ

- 相関がある場合に拡張した一般化最小二乗推定量を導いた
  - $\mathbf{V}$が既知の正定値対称行列であれば,一般化最小二乗推定量はBLUEになる
  - σ^2の不偏推定量も得られた

ここで疑問が一つ

- 実際のデータ解析において, $\mathbf{V}$が既知であることはあるだろうか？
  - $\mathbf{V}$が正しく特定できればBLUEを得られるが, $\mathbf{V}$が間違っていると・・・
- さてどうしたらよいでしょう？