よって、Prop.C14 系より、(i),(ii) が出る。(iii) は、Prop.C15 から出る。
(iv) は $\boldsymbol{e}/\sigma = X$ と考えれば Prop.C18 より分かる。(v) は (iii) より出る。　証明終

[注意]:Prop.1.3.1 より、$\widehat{\boldsymbol{\beta}},\ \boldsymbol{e}$ の平均、分散は知られていたから、正規性が新しい結果である。

Prop.1.3.5.　$\widehat{\boldsymbol{\beta}},\ \widehat{\sigma}^2$ は、$\boldsymbol{\beta},\ \sigma^2$ の不偏推定量の中で分散最小である。
証明：略（レイマンの仮説検定論を見よ）　証明終

モデルの中に含まれている説明変数 $(x_1, \cdots, x_p)$ の一部が不要であるかどうかの検定を考える。

Prop.1.3.6.　$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

$$X = \begin{array}{c} \begin{bmatrix} X_1 & X_2 \end{bmatrix} \\ \begin{array}{cc} p-q & q \end{array} \end{array}, \qquad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} \begin{array}{c} p-q \\ q \end{array}$$

$$H_0 : \boldsymbol{\beta}_2 = \boldsymbol{0}$$

$\Longrightarrow$

(i) H_0 の尤度比 (LR, Likelihood Ratio) は、以下の式で与えられる。

$$LR = \left(||\widetilde{\boldsymbol{e}}\,||^2/||\boldsymbol{e}||^2 \right)^{-\frac{n}{2}}$$

ここに、$\widetilde{\boldsymbol{e}} := \boldsymbol{y} - X_1\widetilde{\boldsymbol{\beta}}_1$

$\widetilde{\boldsymbol{\beta}}_1 : \boldsymbol{\beta}_2 = \boldsymbol{0}$ のときの $\boldsymbol{\beta}_1$ の最小二乗解 (LSE)

(ii) $\widetilde{\boldsymbol{e}} - \boldsymbol{e} \ \sim \ N_n((I_n - \Pi_{X_1})X_2\boldsymbol{\beta}_2,\ \sigma^2(\Pi_X - \Pi_{X_1}))$

$\widetilde{\boldsymbol{y}} := X_1\widetilde{\boldsymbol{\beta}}_1,\ \widetilde{\boldsymbol{e}} - \boldsymbol{e},\ \boldsymbol{e}$ は独立

(iii) H_0 の LR test(有意水準 α) は、

$$\frac{n-p}{q}\frac{||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2}{||\boldsymbol{e}||^2} > F^q_{n-p}(\alpha) \quad \Longleftrightarrow \quad H_0 \text{ を棄てる}$$

証明：(i) Prop.1.3.3 より、

$$\max_{\boldsymbol{\beta},\ \sigma^2} L = K\widehat{\sigma}_M^{-n}\exp(-\tfrac{n}{2})$$

ただし、$\widehat{\sigma}^2_M = \dfrac{1}{n}||\boldsymbol{e}||^2$

H_0 の下で、同様に

$$\max_{\boldsymbol{\beta}_1,\ \sigma^2} L = K\widetilde{\sigma}^{\,-n}\exp(-\tfrac{n}{2})$$

ただし、$\widetilde{\sigma}^{\,2} = \dfrac{1}{n}||\widetilde{\boldsymbol{e}}\,||^2$

ゆえに、$LR = \dfrac{\max_{\boldsymbol{\beta}_1,\ \sigma^2} L}{\max_{\boldsymbol{\beta},\ \sigma^2} L} = \left(\dfrac{\widetilde{\sigma}}{\widehat{\sigma}_M}\right)^{-n} = \left(\dfrac{||\widetilde{\boldsymbol{e}}\,||^2}{||\boldsymbol{e}||^2}\right)^{-\frac{n}{2}}$

(ii) $\boldsymbol{e} = (I - \Pi_X)\boldsymbol{y}$

$$\widetilde{\boldsymbol{e}} - \boldsymbol{e} = (I - \Pi_{X_1})\boldsymbol{y} - (I - \Pi_X)\boldsymbol{y} = (\Pi_X - \Pi_{X_1})\boldsymbol{y}$$

$$\widetilde{\boldsymbol{y}} = \Pi_{X_1}\boldsymbol{y}$$

これらは、1 次変換ゆえ、jointly normal
3 つのベクトルは互いに直交。ゆえに独立となる。
（注：この証明に次頁の命題 A.2 を用いる）

(iii) $LR = \left(\dfrac{||\boldsymbol{e}||^2 + ||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2}{||\boldsymbol{e}||^2}\right)^{-\frac{n}{2}} = \left(1 + \dfrac{||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2}{||\boldsymbol{e}||^2}\right)^{-\frac{n}{2}}$

ゆえに、 $LR < c \iff \dfrac{||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2}{||\boldsymbol{e}||^2} > c'$

H_0 の下で、

$$\left.\begin{array}{r} ||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2 \ \sim \ \sigma^2\chi^2_q \\ ||\boldsymbol{e}||^2 \ \sim \ \sigma^2\chi^2_{n-p} \end{array}\right\} \perp\!\!\!\perp$$

ゆえに、$\dfrac{n-p}{q}\dfrac{||\widetilde{\boldsymbol{e}} - \boldsymbol{e}||^2}{||\boldsymbol{e}||^2} \ \sim \ F^q_{n-p}$ 証明終

この検定を使って、説明変数の選択を行うことができる。
ただし、$x_1, \cdots, x_p$ が重要度の順に並んでいると仮定する。
まず、

$$H_p : \beta_p = 0$$

を検定する。
H_p が棄却されれば $(x_1, \cdots, x_p)$ 全体を使い、採択されれば

$$H_{p-1} : \beta_{p-1} = 0$$

に進む。以下、同様である。この方法を、**変数減少法**という。

逆に、$H_1 : \beta_1 = 0$ から出発し、変数を加えていく方法を **変数増加法** という。
他に、**変数減増法（イナバ氏お勧め）**、**変数増減法** がある。

Remark：変数減増法
3 つ以上の変数（例えば、x_1, x_2, x_3）を削除したとき、本当は x_1 が重要な変数であるのに、これに相関がある x_2, x_3 が残っていることで、最初に、削除されることがある。この後、x_2, x_3 は、単独では影響が小さいから削除されると、本当に残したい x_1 も含めて削除されてしまう。
このような場合に、最適なモデル（今回は、x_1 を含むモデル）に到達できる変数選択法として、変数減増法がある。変数減増法は、フルモデルから始めて、変数を減少していくが、2 つ以上の変数を削除した場合、今までに削除した中で、取り込むと寄与率が著しく上がる時には、モデルに残す、という手法である。大事な変数を落としてしまわない、と言う意味で、広く用いられている選択方式である。

Remark：最小二乗解, LSE(Least Square Estimator) とは
Def.1.2.2 で与えられた $\widehat{\boldsymbol{\beta}}$ は，ズレの 2 乗の和が最小になる解である。これを、**最小二乗解** という。

補題 A.1. 部分行列の逆行列表現の公式

(A.1) $$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1}+A^{-1}BWCA^{-1} & -A^{-1}BW \\ -WCA^{-1} & W \end{pmatrix}$$

が成り立つ。ただし、$W=(D-CA^{-1}B)^{-1}$ である。

定義 A.1. 行列 X による射影 Π_X とは、$\Pi_X = X(X'X)^{-1}X'$ とする。

命題 A.2. $X=[X_1X_2]$ における射影 Π_X, Π_{X_1} について、

(A.2) $$\Pi_{X_1}\Pi_X = \Pi_{X_1}$$

が成り立つことを示せ。

証明：まず、射影の定義 A.1 より、

(A.3) $$\Pi_X = X(X'X)^{-1}X'$$

(A.4) $$\Pi_{X_1} = X_1(X_1'X_1)^{-1}X_1'$$

である。ここで、

(A.5) $$(X'X)^{-1} = \left[\begin{bmatrix} X_1' \\ X_2' \end{bmatrix}[X_1X_2]\right]^{-1} = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}$$

に対して、補題 A.1 より、(A.1) 式を適用することができる。（これで、必ず証明できるはず。）

(A.5)$'$ $$(X'X)^{-1} = \begin{bmatrix} (X_1'X_1)^{-1}+(X_1'X_1)^{-1}X_1'X_2WX_2'X_1(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}X_1'X_2W \\ -WX_2'X_1(X_1'X_1)^{-1} & W \end{bmatrix}$$

$$=: \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$ とおく。

ここで、$W^{-1} = X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2 = X_2'(I-\Pi_{X_1})X_2$ である。

したがって、X の射影

(A.6) $$\Pi_X = X(X'X)^{-1}X' = [X_1X_2]\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} = [X_1X_2]\begin{bmatrix} AX_1'+BX_2' \\ CX_1'+DX_2' \end{bmatrix}$$

$$= X_1AX_1' + X_1BX_2' + X_2CX_1' + X_2DX_2'$$

ここに、(A.5)$'$ を代入して、

(A.7) $$\begin{aligned}\Pi_X &= X_1\{(X_1'X_1)^{-1}+(X_1'X_1)^{-1}X_1'X_2WX_2'X_1(X_1'X_1)^{-1}\}X_1' \\ &\quad +X_1\{-(X_1'X_1)^{-1}X_1'X_2W\}X_2' + X_2\{-WX_2'X_1(X_1'X_1)^{-1}\}X_1' + X_2WX_2' \\ &= \Pi_{X_1} + \Pi_{X_1}X_2WX_2'\Pi_{X_1} - \Pi_{X_1}X_2WX_2' - X_2WX_2'\Pi_{X_1} + X_2WX_2' \\ &= \Pi_{X_1} + (I-\Pi_{X_1})(X_2WX_2')(I-\Pi_{X_1})\end{aligned}$$

　ここで、$X_1'\Pi_{X_1} = X_1'X_1(X_1'X_1)^{-1}X_1' = X_1'$ ゆえ $X_1'(I - \Pi_{X_1}) = 0$ であり、

(A.8)　$X_1'\Pi_X = X_1'\{\Pi_{X_1} + (I - \Pi_{X_1})(X_2WX_2')(I - \Pi_{X_1})\} = X_1' + 0 = X_1'$

となる。したがって、結論を得る。（証明終）

別証明：まず、

(A.9)　$X'\Pi_X = X'X(X'X)^{-1}X' = X'$

ゆえ、$X'\Pi_X = X'$ に $X = [X_1X_2]$ を代入すると、

(A.10)　$\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \Pi_X = \begin{bmatrix} X_1'\Pi_X \\ X_2'\Pi_X \end{bmatrix} = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix}$

　したがって、$X_1'\Pi_X = X_1'$ と分かる。よって、結論を得る。（別証明終）