モデル:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\quad \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix},\quad \mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix},\quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix},\quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$, $\mathbf{X}$ は非確率的

について最小二乗推定量と t 統計量と F 統計量の分布について説明する。

(最小二乗推定量の分布)

最小二乗推定量は

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

である。これに $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ を代入すると

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon} \\ &= \boldsymbol{\beta} + (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon} \end{aligned}$$

となる。誤差項は $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$ であり $\mathbf{X}$ は非確率的、さらに $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon}$ は $\boldsymbol{\varepsilon}$ の 1 次変換であるので, $\hat{\boldsymbol{\beta}}$ は多変量正規分布に従いその平均は

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}E(\boldsymbol{\varepsilon}) = \boldsymbol{\beta}$$

分散は

$$\begin{aligned} Var(\hat{\boldsymbol{\beta}}) &= E[\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^{\mathrm{T}}\,] \\ &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}})\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}, \end{aligned}$$

となる。つまり

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1})$$

である。

(カイ 2 乗分布)

標準正規分布の 2 乗の和の分布は「カイ 2 乗分布」と呼ばれる。より正確には $i = 1,\ldots,m$ について X_i は互いに独立に $X_i \sim N(0,1)$ であるとすると、この m 個の標準正規確率変数の 2 乗の和:

$$Y = X_1^2 + X_2^2 + \cdots + X_m^2$$

の分布は「自由度 m のカイ二乗分布」と言われ $Y \sim \chi^2(m)$ と表記される。

(t 分布)

確率変数 Z と Y は互いに独立で $Z \sim N(0,1)$ および $Y \sim \chi^2(m)$ であるとする。この時

$$T = Z / \sqrt{(Y/m)}$$

で定義される確率変数 T の分布は「自由度 m の t 分布」と呼ばれ、$T \sim t(m)$ と表記される。

(t 統計量の定義)

$\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)^{\mathrm{T}}$とする。$\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$ は k 行 k 列の行列である、その (i, j) 成分は$\hat{\beta}_i$ と $\hat{\beta}_j$ の(共)分散である。k 行 k 列の行列 $(\mathbf{X}^T\mathbf{X})^{-1}$の (i, j) 成分を q_{ij} と表す事にする。$\hat{\beta}_i$ は正規分布に従いその平均は β_i 、分散は$\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$の$(i, i)$成分 $\sigma^2 q_{ii}$ である。つまり$\hat{\beta}_i \sim N(\beta_i, \sigma^2 q_{ii})$ である。よって、$\hat{\beta}_i$ 基準化したものを Z とすると、Z は

$$Z = \frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{q_{ii}}} \sim N(0, 1)$$

となる。よってもし σ^2 が既知であれば、この Z を検定統計量としてβ_iについての仮説検定を行うことができる。例えば帰無仮説 $H_0: \beta_i = 0$ の検定には

$$Z = \frac{\hat{\beta}_i}{\sigma\sqrt{q_{ii}}}$$

を用いる。これは上記の結果より、帰無仮説のもとで $N(0, 1)$ に従うので$|Z| > 1.96$であれば、有意水準$\alpha = 0.05$で帰無仮説を棄却する。

しかし実際にはσは未知であるので、それをσ^2の OLS 推定量:

$$s^2 = \frac{1}{n-k}\sum_{i=1}^{n} e_i^2 = \frac{1}{n-k}\mathbf{e}^{\mathrm{T}}\mathbf{e}, \quad \text{ここで}\ \mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

の平方根で置き換える。これが t 統計量であり

$$t_n = \frac{\hat{\beta}_i - \beta_i}{s\sqrt{q_{ii}}}$$

と定義される。これを用いてβ_i についての仮説検定を行う

(t 統計量の分布)

t 統計量は

$$t_n = \frac{\hat{\beta}_i - \beta_i}{s\sqrt{q_{ii}}} = \frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{q_{ii}}}\frac{\sigma}{s} = \frac{Z}{s/\sigma} = \frac{Z}{\sqrt{s^2/\sigma^2}} = \frac{Z}{\sqrt{\dfrac{(n-k)s^2}{\sigma^2}\dfrac{1}{n-k}}} = \frac{Z}{\sqrt{Y/(n-k)}}$$

と書く事ができる。ここで$Z = \dfrac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{q_{ii}}}$、$Y = \dfrac{(n-k)s^2}{\sigma^2}$である。先ほどの議論により$Z \sim N(0, 1)$ であるのでもし Y が $\chi^2(n-k)$ に従えば(さらに Z と Y が独立であれば)、$t_n \sim t\ (n-k)$ となる事がわかる。以下では Y が $\chi^2(n-k)$ に従う事を示す。

残差 $\mathbf{e}$ を書き直すと

$$\begin{aligned}\mathbf{e} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} - \mathbf{X}\hat{\boldsymbol{\beta}} = -\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \boldsymbol{\varepsilon} &= -\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon} \\ &= [\mathbf{I}_n - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}]\boldsymbol{\varepsilon} \\ &= \mathbf{M}\boldsymbol{\varepsilon},\end{aligned}$$

である。ここで $\mathbf{M} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ である。$\mathbf{M}$ は直接計算する事により

$$\mathbf{M} = \mathbf{M}^{\mathrm{T}} = \mathbf{M}^2$$

という性質を持っている事がわかる。 ある行列 $\mathbf{A}$ が $\mathbf{A} = \mathbf{A}^2$ を満たす時、その行列はベキ等行

列と呼ばれる。$\mathbf{M}$ は対称なベキ等行列である。以下の事実が重要である。

[ベキ等行列による正規分布の 2 次形式の分布]
「$\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_n)$ で $\mathbf{A}$ を対称なベキ等行列でありそのランクは $\mathrm{rank}(\mathbf{A}) = r$ であるとする。この時 $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \sim \chi^2(r)$ である。」

行列 $\mathbf{M}$ はベキ等であるからそのランクはトレースに等しく $\mathrm{rank}(\mathbf{M}) = \mathrm{tr}(\mathbf{M}) = n - k$ である。よって $\boldsymbol{\varepsilon}/\sigma \sim N(0,\ \mathbf{I}_n)$ である事に注意すると

$$(n-k)\frac{s^2}{\sigma^2} = \frac{1}{\sigma^2}\mathbf{e}^{\mathrm{T}}\mathbf{e} = \frac{1}{\sigma^2}\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{M}\boldsymbol{\varepsilon} = \left(\frac{\boldsymbol{\varepsilon}}{\sigma}\right)^{\mathrm{T}}\mathbf{M}\left(\frac{\boldsymbol{\varepsilon}}{\sigma}\right) \sim \chi^2(n-k)$$

を得る。よって Y は $\chi^2(n-k)$ に従う。さらに Z と Y が独立である事も示すことができるので(付録参照) $t_n \sim t\ (n-k)$ となる事がわかる。

(F 分布)
2 つの確率変数 X_1 と X_2 がそれぞれ独立に $X_1 \sim \chi^2(m)$ と $X_2 \sim \chi^2(n)$ であるとする。この時

$$F = \frac{X_1/m}{X_2/n}$$

と定義される確率変数 F の分布を「自由度 m, n の F 分布」と呼び、$F \sim F(m, n)$と表記する。

(F 統計量の定義)
F 統計量とは一般的に線形制約 H_0: $\mathbf{C}\boldsymbol{\beta} = \mathbf{r}$ を検定するための統計量である。ここで $\mathbf{C}$ は q 行 k 列の行列、$\mathbf{r}$ は q 行 1 列のベクトルである。例えば $\mathbf{C} = \mathbf{I}_k$ $(q = k)$ であれば、この制約は

$$\boldsymbol{\beta} = \mathbf{r} \quad \text{つまり} \quad \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}$$

という係数ベクトルががあるベクトル $\mathbf{r}$ に等しいという仮説になり、またさらに $\mathbf{r} = \mathbf{0}$ ($\mathbf{0}$ は成分が全て 0 のベクトル)であれば、係数ベクトルが全て 0 であるという仮説になる。

F 検定統計量は一般に

$$F = (\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})^{\mathrm{T}}[s^2\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})/q = \frac{(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})^{\mathrm{T}}[\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})/q}{s^2}$$

と定義され、帰無仮説 H_0: $\mathbf{C}\boldsymbol{\beta} = \mathbf{r}$ のもとで $F \sim F(q, n-k)$ となる。

(F 統計量の分布)
まず $\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r}$ の分布を求める。$\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{C}\boldsymbol{\beta} = \mathbf{C}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ であり、またこれまでの議論より

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \sim N(\mathbf{0}, \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1})$$

であるから $\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{C}\boldsymbol{\beta} \sim N(\mathbf{0},\ \sigma^2\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}})$ である。さらに帰無仮説のもとでは $\mathbf{C}\boldsymbol{\beta} = \mathbf{r}$ なので、

帰無仮説のもとでは

$$\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r} \sim N(\mathbf{0},\ \sigma^2 \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}})$$

となる。以下の事実が重要である。

[正規分布の分散の逆行列による 2 次形式の分布]

「$\mathbf{x}$ は q 行 1 列であり、$\mathbf{x} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathrm{rank}(\boldsymbol{\Sigma}) = q$ である時 $\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} \sim \chi^2(q)$ である。」

これよりすぐさま

$$(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})^{\mathrm{T}}[\sigma^2 \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r}) \sim \chi^2(q)$$

が導かれる。ここで σ^2 は未知なので、σ^2 をその OLS 推定量の s^2 で置き換え、さらに q で割ったものが F 統計量である。未知パラメータを推定量で置き換えたので F 統計量は $\chi^2(q)$ には従わない。以下 F 統計量の分布を求めよう。

既に見たように

$$(n-k)\frac{s^2}{\sigma^2} \sim \chi^2(n-k)$$

である。F 統計量を

$$F = \frac{(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})^{\mathrm{T}}[\sigma^2 \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})/q}{s^2/\sigma^2} = \frac{X_1/q}{X_2/(n-k)}$$

ここで

$$X_1 = (\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r})^{\mathrm{T}}[\sigma^2 \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r}),\quad X_2 = (n-k)\frac{s^2}{\sigma^2}$$

と書き直すと今までの議論より $X_1 \sim \chi^2(q)$、$X_2 \sim \chi^2(n-k)$ である事がわかる。さらにこの 2 つは独立である事を示すことができるので(付録参照)、結局 $F \sim F(q, n-k)$ である事がわかる。

(R の出力について)

R で lm 関数を使って回帰分析を行うと分析結果として

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.230e+03  6.120e+02  -3.644 0.004507 **
X1           3.728e-01  6.059e-02   6.153  0.000108 ***
X2           4.635e-01  8.621e-02   5.376  0.000312 ***
X3           1.142e+02  2.509e+01   4.549  0.001059 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 404.1 on 10 degrees of freedom
Multiple R-squared: 0.9804,    Adjusted R-squared: 0.9745
F-statistic: 166.7 on 3 and 10 DF,  p-value: 7.765e-09
```

のように出力される。ここでの t value というのが t 統計量の $H_0: \beta_i = 0$ のもとでの値であり、Pr(>|t|) は t 分布に従う確率変数がこの t 統計量の絶対値より大きくなる確率である。またこの F-statistic というのは(定数項を除く)説明変数の係数が全て 0 であるという帰無仮説のための F 統計量の値であり、p-value は F 分布に従う(上記の例では自由度 3 と10 の F 分布)確率変数がこの F 統計量の値より大きくなる確率である。

(付録： 証明に興味がある人だけ読んでください)
ここでは前述した
(1) [ベキ等行列による正規分布の 2 次形式の分布]
「$\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_n)$ で $\mathbf{A}$ を対称なベキ等行列でありそのランクは $\mathrm{rank}(\mathbf{A}) = r$ であるとする。この時 $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \sim \chi^2(r)$ である。」事、

(2) t 統計量の分子と分母が独立である事、

(3) [正規分布の分散の逆行列による 2 次形式の分布]
「$\mathbf{x}$ は q 行 1 列で、$\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Sigma})$, $\mathrm{rank}(\mathbf{\Sigma}) = q$ である時 $\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x} \sim \chi^2(q)$ である。」事

および

(4) F 統計量の分子と分母が独立である事

を証明する。

以下では行列の固有値、行列のランク(階数)および、実対称行列の直行行列による対角化についての知識がある事を前提とする。

[(1)の証明]
ベキ等行列 $\mathbf{A}$ の固有値は 0 か 1 であり、$\mathrm{rank}\ (\mathbf{A}) = r$ であるから、対称行列 $\mathbf{A}$ はある直行行列 $\mathbf{P}$ によって

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

と対角化する事ができる。ここで $\mathbf{0}$ は要素が全て 0 の行列である。この $\mathbf{P}$ を用いて $\mathbf{z}$ を

$$\mathbf{z} = \mathbf{P}^{\mathrm{T}}\mathbf{x}$$

と定義すると

$$E(\mathbf{z}) = \mathbf{P}^{\mathrm{T}}E(\mathbf{x}) = 0 \quad \text{および} \quad Var(\mathbf{z}) = E(\mathbf{z}\mathbf{z}^{\mathrm{T}}) = \mathbf{P}^{\mathrm{T}}E(\mathbf{x}\mathbf{x}^{\mathrm{T}})\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{I}_n\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}_n$$

(直行行列の $\mathbf{P}\mathbf{P}^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}}\mathbf{P} = \mathbf{I}_n$ という性質を使っている)であるから $\mathbf{z} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_n)$ である。また $\mathbf{x}$ は $\mathbf{z}$ を使って $\mathbf{x} = \mathbf{P}\mathbf{z}$ と表すことができる事に注意すると

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{z}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}\mathbf{z} = \mathbf{z}^{\mathrm{T}}\begin{bmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{bmatrix}\mathbf{z} = \sum_{i=1}^{r} z_i^2$$

となる。ここで $z_i \sim N(0, 1)$ であるから、$\sum_{i=1}^{r} z_i^2 \sim \chi^2(r)$ となる。

[(2)の証明]
次の事実を証明なしで述べる。

[正規分布に従う確率変数の１次関数と 2 次形式の独立]
「$\mathbf{x} \sim N(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$, $\mathbf{B}$ は q 行 n 列の行列、$\mathbf{A}$ は n 行 n 列の対称行列とする。この時 $\mathbf{B}\mathbf{A} = \mathbf{0}$ であれば $\mathbf{B}\mathbf{x}$ と $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ は独立である。」

これを用いて証明する。

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon}, \quad \text{および} \quad \mathbf{e}^{\mathrm{T}}\mathbf{e} = \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{M}\boldsymbol{\varepsilon}$$

であり、さらに

$$
\begin{aligned}
(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{M} &= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}) \\
&= (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}} - (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}} \\
&= \mathbf{0}
\end{aligned}
$$

であるので、上の事実より $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}$ と $\mathbf{e}^{\mathrm{T}}\mathbf{e}$ は独立である。よって $\hat{\beta}_i - \beta_i$ と s^2 は独立である。

[(3)の証明]

まず $\mathbf{\Sigma}$ はある正則な行列 $\mathbf{C}$ によって

$$\mathbf{C}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{C} = \mathbf{I}_q$$

となる事を示す。$\mathbf{\Sigma}$ は対称行列なので

$$\mathbf{P}^{\mathrm{T}}\mathbf{\Sigma}\,\mathbf{P} = \mathbf{D},\quad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_q \end{bmatrix}$$

となる直行行列 $\mathbf{P}$ が存在する。ここで $\lambda_1, \ldots, \lambda_q$ は $\mathbf{\Sigma}$ の固有値であり、仮定よりすべて 0 ではない ($\mathbf{\Sigma}$ は仮定より正値定符号行列なのでこれはすべて正である事に注意)。ここで

$$\mathbf{D}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_q} \end{bmatrix}$$

とすれば $\mathbf{D} = \mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}$ でありこの $\mathbf{D}^{\frac{1}{2}}$ の逆行列を用いると

$$\mathbf{P}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{P} = \mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}} \Leftrightarrow \mathbf{D}^{-\frac{1}{2}}\mathbf{P}^{\mathrm{T}}\mathbf{\Sigma}\;\mathbf{P}\;\mathbf{D}^{-\frac{1}{2}} = \mathbf{I}_q$$

であるので $\mathbf{C} = \mathbf{P}\mathbf{D}^{-\frac{1}{2}}$ が求める $\mathbf{C}$ である。

この $\mathbf{C}$ によって $\mathbf{\Sigma}$ は $\mathbf{\Sigma} = (\mathbf{C}^{\mathrm{T}})^{-1}\mathbf{C}^{-1} = (\mathbf{C}\,\mathbf{C}^{\mathrm{T}})^{-1}$ と表せられるので、$\mathbf{\Sigma}$ の逆行列は

$$\mathbf{\Sigma} = \mathbf{C}\,\mathbf{C}^{\mathrm{T}}$$

となる。よって $\mathbf{z} = \mathbf{C}^{\mathrm{T}}\mathbf{x}$ とおくと

$$\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{C}\,\mathbf{C}^{\mathrm{T}}\mathbf{x} = (\mathbf{C}^{\mathrm{T}}\mathbf{x})^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{x} = \mathbf{z}^{\mathrm{T}}\mathbf{z} = \sum_{i=1}^{q} z_i^2$$

となる。ここで $\mathbf{z}$ は正規分布であり

$$E(\mathbf{z}) = E(\mathbf{C}^{\mathrm{T}}\mathbf{x}) = \mathbf{C}^{\mathrm{T}}E(\mathbf{x}) = \mathbf{0},\quad Var(\mathbf{z}) = E(\mathbf{z}\mathbf{z}^{\mathrm{T}}) = \mathbf{C}^{\mathrm{T}}E(\mathbf{x}\mathbf{x}^{\mathrm{T}})\mathbf{C} = \mathbf{C}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{C} = \mathbf{I}_q$$

であるので

$$z_i \sim N(0, 1)$$

となり、よって $\displaystyle\sum_{i=1}^{q} z_i^2 \sim \chi^2(q)$ となる。

[(4)の証明]

次の事実を証明なしで述べる。

[正規分布に従う確率変数の1次関数と 2 次形式の独立]

「$\mathbf{x} \sim N(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$, $\mathbf{A}$ と $\mathbf{B}$ はともに n 行 n 列の対称行列とする。この時 $\mathbf{AB} = \mathbf{0}$ であれば $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ と $\mathbf{x}^{\mathrm{T}}\mathbf{B}\mathbf{x}$ は独立である。」

H_0: $\mathbf{C}\boldsymbol{\beta} = \mathbf{r}$ が正しいとき

$$\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{r} = \mathbf{C}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\varepsilon} = \mathbf{G}\boldsymbol{\varepsilon}$$

ここで $\mathbf{G} = \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ である。よって F 統計量のところで出てきた X_1 は

$$\begin{aligned} X_1 &= \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}[\sigma^2\,\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}\mathbf{G}\boldsymbol{\varepsilon} \\ &= \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{A}\boldsymbol{\varepsilon} \end{aligned}$$

と表すことができる。ここで $\mathbf{A}= \mathbf{G}^{\mathrm{T}}[\sigma^2\,\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}\mathbf{G}$ である。X_2 は

$$(n-k)s^2 = \mathbf{e}^{\mathrm{T}}\mathbf{e} = \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{M}\boldsymbol{\varepsilon}$$

であるので

$$X_2 = \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{B}\boldsymbol{\varepsilon}$$

と表すことができる。ここで $\mathbf{B} = \mathbf{M}/\sigma^2$ である。先ほどの事実より $\mathbf{AB} = \mathbf{0}$ であれば X_1 と X_2 は独立である。これは

$$\mathbf{AM} = \mathbf{G}^{\mathrm{T}}[\sigma^2\,\mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{C}^{\mathrm{T}}]^{-1}\mathbf{GM}$$

であるが

$$\mathbf{GM} = \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}[\mathbf{I}_n - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}] = \mathbf{C}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\,[\mathbf{X}^{\mathrm{T}} - \mathbf{X}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}] = \mathbf{0}$$

であるので $\mathbf{AM} = \mathbf{0}$ となる事より示すことができる。