第1章　統計学のレビュー

1.1　標本分布

定義 1.1（χ^2 分布）

$Z_i \overset{i.i.d.}{\sim} N(0,1),\ i=1,\cdots,n$ のとき，$X=\sum_{i=1}^n Z_i^2$ の従う確率分布を**自由度 n の χ^2 分布**といい，このことを，$X \sim \chi^2(n)$ で表すことにする．

定理 1.1（χ^2 分布の確率密度関数）

自由度 n の χ^2 分布の確率密度関数は次式で与えられる．

$$f(x)=\frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}},\quad x>0 \tag{1.1}$$

ただし，ここで，Γ は,

$$\Gamma(x)=\int_0^\infty \mathrm{e}^{-t}t^{x-1}\mathrm{d}t,\quad x>0$$

で定義される**ガンマ関数**である [1]．

[1] ガンマ関数については,

$$\Gamma(n+1)=n!,\quad n\in\mathbb{N},$$
$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$$

が成立する．

証明　岩城（2012）問題 10.2 参照．　□

定理 1.2（χ^2 分布の再生性）

X_1 と X_2 が互いに独立で，$X_i \sim \chi^2(n_i),\ i=1,2$ とすると,

$$X_1+X_2 \sim \chi^2(n_1+n_2).$$

証明　$X_i \sim \chi^2(n_i)$ の確率密度関数を $f_{n_i},\ i=1,2$ としたとき，X_1 と X_2 が独立であることから，X_1+X_2 の確率密度関数は,

$$f(x)=\int_0^x f_{n_1}(x-y)f_{n_2}(y)\mathrm{d}y,\quad x>0$$

で与えられる（岩城（2012）定理 10.3 参照.）．したがって，上式が，$n=n_1+n_2$ として，(1.1) となることを示せばよい．(1.1) より,

$$f(x)=\int_0^x \frac{2^{-\frac{n_1}{2}}}{\Gamma(\frac{n_1}{2})}(x-y)^{\frac{n_1}{2}-1}\mathrm{e}^{-\frac{(x-y)}{2}}\times\frac{2^{-\frac{n_2}{2}}}{\Gamma(\frac{n_2}{2})}y^{\frac{n_2}{2}-1}\mathrm{e}^{-\frac{y}{2}}\mathrm{d}y$$

$$= \frac{2^{-\frac{n_1+n_2}{2}}}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} \mathrm{e}^{-\frac{x}{2}} \int_0^x (x-y)^{\frac{n_1}{2}-1} y^{\frac{n_2}{2}-1} \mathrm{d}y.$$

ここで，$t = \frac{y}{x}$ として変数置換すると，最右辺の積分は，

$$\int_0^1 (1-t)^{\frac{n_1}{2}-1} t^{\frac{n_2}{2}-1} x^{\frac{n_1+n_2}{2}-1} \mathrm{d}t = x^{\frac{n_1+n_2}{2}-1} \int_0^1 (1-t)^{\frac{n_1}{2}-1} t^{\frac{n_2}{2}-1} \mathrm{d}t.$$

右辺の積分は，ベータ関数の定義 [2] により，独立変数の値が $\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$ のベータ関数 $B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$ であり，ベータ関数の性質 [3] から，$B\left(\frac{n_1}{2}, \frac{n_2}{2}\right) = \frac{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}+\frac{n_2}{2}\right)}$ が成立する．以上により，

$$f(x) = \frac{2^{-\frac{n_1+n_2}{2}}}{\Gamma(\frac{n_1+n_2}{2})} x^{\frac{n_1+n_2}{2}-1} \mathrm{e}^{-\frac{x}{2}}, \quad x > 0$$

となり，題意が成立する．　□

定理 1.3（標本平均・分散の分布）

$X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2),\ i = 1, \cdots, n$ のとき，

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}, \quad \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \tag{1.2}$$

は互いに独立で，各々，$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$，$\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$ となる．

証明　n 個の n 次ベクトル $\boldsymbol{t}_i,\ i = 1, \cdots, n$ を

$$\begin{aligned}
\boldsymbol{t}_1 &= \left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}\right), \\
\boldsymbol{t}_2 &= \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, \cdots, 0\right), \\
\boldsymbol{t}_3 &= \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, 0, \cdots 0\right), \\
&\ \vdots \\
\boldsymbol{t}_n &= \left(\frac{1}{\sqrt{n(n-1)}}, \cdots, \frac{1}{\sqrt{n(n-1)}}, -\frac{n-1}{\sqrt{n(n-1)}}\right)
\end{aligned}$$

として，$n \times n$ 行列 $\boldsymbol{T}$ を

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{t}_1 \\ \boldsymbol{t}_2 \\ \boldsymbol{t}_3 \\ \vdots \\ \boldsymbol{t}_n \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & -\frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix},$$

とする．また，

2)

$$B(x,y) = \int_0^1 (1-t)^{x-1} t^{y-1} \mathrm{d}t, \quad x, y > 0$$

で定義される 2 変数関数を**ベータ関数**という．

3)　ベータ関数とガンマ関数の間には次式が成立する．

$$B(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}.$$

$$\boldsymbol{X}-\boldsymbol{\mu}=\begin{pmatrix} X_1-\mu \\ X_2-\mu \\ \vdots \\ x_n-\mu \end{pmatrix}$$

として，

$$\boldsymbol{Y}=\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}=\boldsymbol{T}(\boldsymbol{X}-\boldsymbol{\mu})=\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i-\mu) \\ \frac{1}{\sqrt{2}}(X_1-X_2) \\ \frac{1}{\sqrt{6}}(X_1+X_2-2X_3) \\ \vdots \\ \frac{1}{\sqrt{n(n-1)}}(\sum_{i=1}^{n-1}X_1-(n-1)X_n) \end{pmatrix}$$

とおく．

はじめに，$Y_i \overset{i.i.d.}{\sim} N(0,\sigma^2),\, i=1,\cdots,n$ を示す．

$X_i \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$，かつ，$Y_i$ は，$(X_1,\cdots,X_n)$ の線形変換であるから，正規分布に従う．ここで，$\boldsymbol{T}$ は直交行列，すなわち，$\boldsymbol{I}$ を n 次単位行列として，

$$\boldsymbol{T}\boldsymbol{T}^\top=\boldsymbol{T}^\top\boldsymbol{T}=\boldsymbol{I}$$

であることに注意すると，$X_i \overset{i.i.d.}{\sim} N(\mu,\sigma^2),\, i=1,\cdots,n$ であるから，

$$\begin{aligned}
\mathbb{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^\top\right]&=\sigma^2\boldsymbol{I}, \\
\mathbb{E}[Y_i]&=\mathbb{E}[\boldsymbol{t}_i(\boldsymbol{X}-\boldsymbol{\mu})]=\boldsymbol{t}_i(\mathbb{E}[\boldsymbol{X}]-\boldsymbol{\mu})=0, \\
\mathbb{E}[Y_iY_j]&=\mathbb{E}\left[\boldsymbol{t}_i(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^\top\boldsymbol{t}_j^\top\right], \\
&=\boldsymbol{t}_i\mathbb{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^\top\right]\boldsymbol{t}_j^\top=\sigma^2\boldsymbol{t}_i\boldsymbol{t}_j^\top=\sigma^2\delta_{ij}, \\
\sigma^2\delta_{ij}&=\begin{cases} 1 & i=j, \\ 0 & i\neq j \end{cases}.
\end{aligned}$$

$\therefore\ Y_i \overset{i.i.d.}{\sim} N(0,\sigma^2),\, i=1,\cdots,n$ となる．

次に，$\bar{X}\sim N\left(\mu,\sigma^2/n\right)$ を示す．これは，

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu+\mu)=\frac{1}{\sqrt{n}}Y_1+\mu \tag{1.3}$$

より，明らかである．

最後に，$\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2\sim\chi^2(n-1)$，かつ，これが $\bar{X}$ と独立であることを示す．

$\boldsymbol{T}$ は直交行列であったことに注意すると，

$$\boldsymbol{Y}^\top\boldsymbol{Y}=(\boldsymbol{X}-\boldsymbol{\mu})^\top\boldsymbol{T}^\top\boldsymbol{T}(\boldsymbol{X}-\boldsymbol{\mu})=(\boldsymbol{X}-\boldsymbol{\mu})^\top(\boldsymbol{X}-\boldsymbol{\mu})$$

すなわち，$\sum_{i=1}^{n}(X_i-\mu)^2=\sum_{i=1}^{n}Y_i^2$．ここで，(1.3) より，$Y_1=\sqrt{n}\left(\bar{X}-\mu\right)$

に注意すると，

$$
\begin{aligned}
\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2 &= \sum_{i=1}^{n}\left\{(X_i-\mu)-(\bar{X}-\mu)\right\}^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\sum_{i=1}^{n}(X_i-\mu)(\bar{X}-\mu) + \sum_{i=1}^{n}(\bar{X}-\mu)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - 2n(\bar{X}-\mu)^2 + n(\bar{X}-\mu)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2 - Y_1^2 = \sum_{i=2}^{n} Y_i^2.
\end{aligned}
$$

$\therefore \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 = \sum_{i=2}^{n}\left(\frac{Y_i}{\sigma}\right)^2$，かつ，$Y_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ であったから，$\bar{X}$ と $\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2$ は独立であり，かつ，$\frac{Y_i}{\sigma} \overset{i.i.d.}{\sim} N(0,1)$ であるから，定義 1.1 により，$\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$. □

定義 1.2（t 分布）

Z と Y は互いに独立で，各々 $Z \sim N(0,1)$, $Y \sim \chi_n^2$ とする．このとき，$X = \frac{Z}{\sqrt{Y/n}}$ の従う確率分布を**自由度** n **の t 分布**といい，このことを，$X \sim \mathrm{t}(n)$ で表すことにする．

定理 1.4（t 分布確率密度関数）

自由度 n の t 分布の確率密度関数 $f(x,n)$ は次式で与えられる．

$$
f(x;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1+\frac{x^2}{2}\right)^{\frac{n+1}{2}}}, \quad t \in \mathbb{R}. \tag{1.4}
$$

証明　$U \sim N(0,1)$，$V \sim \chi^2(n)$ として，U, V が互いに独立であるとすると，(U,V) の同時密度関数 $g(u,v)$ は，(1.1) より，

$$
g(u,v) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{u^2}{2}} \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} v^{\frac{n}{2}-1}\mathrm{e}^{-\frac{v}{2}}, \quad (u,v) \in \mathbb{R}\times\mathbb{R}_+.
$$

ここで $x = \frac{u}{\sqrt{v/n}}$，$y = v$ として，変数置換すると，ヤコビアンが $|J| = \sqrt{\frac{y}{n}}$ となるので，$(X,Y) = (\frac{U}{\sqrt{V/n}}, V)$ の同時確率密度関数 $h(x,y)$ は,

$$
\begin{aligned}
h(x,y) &= g\left(\frac{x\sqrt{y}}{\sqrt{n}}, y\right)|J| \\
&= \frac{1}{\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} y^{\frac{n}{2}-1} \exp\left(-\frac{y}{2}\left(1+\frac{x^2}{n}\right)\right)\sqrt{\frac{y}{n}}.
\end{aligned}
$$

T の周辺確率密度関数は,

$$
\begin{aligned}
f(x) &= \int_0^\infty h(x,y)\mathrm{d}y \\
&= \int_0^\infty \frac{1}{\sqrt{2n\pi}\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}} y^{\frac{n+1}{2}-1}\exp\left(-\frac{y}{2}\left(1+\frac{x^2}{n}\right)\right)\mathrm{d}y.
\end{aligned}
$$

さらにここで，$z=\frac{y}{2}\left(1+\frac{x^2}{n}\right)$ と変数置換すると，

$$
\begin{aligned}
f(x) &= \int_0^\infty \frac{1}{\sqrt{2n\pi}\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}\left(\frac{2z}{1+\frac{x^2}{n}}\right)^{\frac{n+1}{2}-1}\mathrm{e}^{-z}\left(\frac{2}{1+\frac{x^2}{n}}\right)\mathrm{d}z \\
&= \frac{1}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)\left(1+\frac{x^2}{2}\right)^{\frac{n+1}{2}}}\int_0^\infty z^{\frac{n+1}{2}-1}\mathrm{e}^{-z}\mathrm{d}z \\
&= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\frac{1}{\left(1+\frac{x^2}{2}\right)^{\frac{n+1}{2}}}, \quad x\in\mathbb{R}.
\end{aligned}
$$

∴題意を得る．　□

定理 1.5（標本平均の分布:分散未知の場合）

$X_i \overset{i.i.d}{\sim} N(\mu,\sigma^2),\ i=1,\cdots,n$ のとき，

$$
\bar{X}=\frac{\sum_{i=1}^n X_i}{n},\quad S^2=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\bar{X}\right)^2
$$

とすると，

$$
T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim \mathrm{t}(n).
$$

証明　定理 1.3 より，$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\sim N(0,1)$, $\frac{S^2(n-1)}{\sigma^2}\sim\chi^2(n-1)$ である．

$$
\therefore\qquad T=\frac{\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}}{\left(\frac{S\sqrt{n-1}}{\sigma}\right)/\sqrt{n-1}}\sim \mathrm{t}(n-1).
$$

□

定理 1.6（標本平均の差の分布:等分散の場合）

$X_{i_j} \sim N(\mu_i, \sigma^2), i_j = 1, \cdots, n_i, i = 1, 2$ とし, $X_{i_j}, i_j = 1, \cdots, n_i, i = 1, 2$ は互いに独立であるとする．このとき,

$$\bar{X}_i = \frac{1}{n_i} \sum_{i_j=1}^{n_i} X_{i_j},$$
$$S_i^2 = \sum_{i_j=1}^{n_i} \frac{(X_{i_j} - \bar{X}_i)^2}{n_i - 1}, i = 1, 2,$$
$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$ [4)]

とすると,

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \mathrm{t}(n_1 + n_2 - 2).$$

4) S_p は S_1 と S_2 の加重平均であり, σ^2 の不偏推定量になっている.

証明　$X_{i_j} \sim N(\mu_i, \sigma^2)$, $i_j = 1, \cdots, n_i$, $i = 1, 2$ かつ, X_{i_j}, $i_j = 1, \cdots, n_i$, $i = 1, 2$ は互いに独立であるから, $\bar{X}_i \sim N(\mu_i, \sigma^2/n_i)$, $i = 1, 2$ であり,

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0, 1).$$

一方, $(n_i - 1)S_i^2/\sigma^2 \sim \chi^2(n_i - 1)$, $i = 1, 2$ であり, それらは, 互いに独立である. $\therefore$ χ^2 分布の再生性 (定理 1.2) により, $(n_1 + n_2 - 2)S_p^2/\sigma^2 \sim \chi^2(n_1 + n_2 - 2)$. よって,

$$T = \frac{\{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)\}/\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}{\sqrt{(n_1 + n_2 - 2)S_p^2/\{(n_1 + n_2 - 2)\sigma^2\}}}$$

であるから, t 分布の定義（定義 1.2）により, 題意を得る. □

1.2　線形回帰

1.2.1　モデル

$(y_t, x_{1t}, \cdots, x_{kt})$, $t = 1, \cdots, T$ を所与として, 次式を満たす未知の係数 $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k)^\top$ を推定する問題を考える.

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + \epsilon_t, \quad t = 1, \cdots, T \tag{1.5}$$

ただし, ここで, $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_T)^\top$ は, **撹乱項**で次の仮定を満たすものとする.

仮定 1.1 (線形回帰式：撹乱項の仮定)**.**

(i) $\mathbb{E}[\epsilon_t] = 0 \quad \forall t.$

(ii) $\mathbb{E}[\epsilon_s \epsilon_t] = \begin{cases} \sigma^2 & s = t, \\ 0 & s \neq t \end{cases} \quad \forall s, t.$

(iii) $\epsilon_t \sim N(0, \sigma^2), \quad \forall t.$

この仮定を，ベクトル表現すれば，

$$\boldsymbol{\epsilon} \sim N_T(\mathbf{0}_T, \sigma^2 \boldsymbol{I}_T). \tag{1.6}$$

ただし，ここで，$\mathbf{0}_T$ は T 次のゼロベクトル．$\boldsymbol{I}_T$ は $T \times T$ 単位行列で，$N_T(\mathbf{0}_T, \sigma^2 \boldsymbol{I}_T)$ は，平均ベクトル $\mathbf{0}_T$，分散共分散行列 $\sigma^2 \boldsymbol{I}_T$ の T 次元正規分布を表す．また，(1.5) を行列表現すると，

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1.7}$$

$$\Longleftrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{k1} \\ & \vdots & \\ x_{1T} & \cdots & x_{kT} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_T \end{pmatrix} \tag{1.8}$$

となる．以下では，$\mathrm{rank}(\boldsymbol{X}) = k$ とする．

定義 1.3（最小２乗推定量）

$$\min_{\boldsymbol{\beta}} \sum_{t=1}^{T} (y_t - x_{1t}\beta_1 - \cdots - x_{kt}\beta_k)^2 = \min_{\boldsymbol{\beta}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

となるように，$\boldsymbol{\beta}$ を推定する方法を**最小２乗法**という．最小２乗法によって求められた推定量を**最小２乗推定量**という．

いまの場合，最小化の一階条件：

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) &= \frac{\partial}{\partial \boldsymbol{\beta}} (\boldsymbol{y}^\top \boldsymbol{y} - 2\boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{y} + \boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta}) \\ &= -2(\boldsymbol{X}^\top \boldsymbol{y} - \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta}) = \mathbf{0} \end{aligned} \tag{1.9}$$

より，$\boldsymbol{\beta}$ の最小 2 乗推定量 $\hat{\boldsymbol{\beta}}$ は，

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \boldsymbol{y} \tag{1.10}$$

となる．

定義 1.4（残差）

(i) $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$ を**理論値**あるいは**予測値**といい，
(ii) $\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$ を**残差**という．

注 1.1. (1.9) より，

$$\boldsymbol{X}^\top \boldsymbol{e} = \boldsymbol{X}^\top (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) = \mathbf{0}.$$

さらに，これより，

$$\hat{\boldsymbol{y}}^\top \boldsymbol{e} = \boldsymbol{e}^\top \hat{\boldsymbol{y}} = (\hat{\boldsymbol{\beta}}^\top \boldsymbol{X}^\top \boldsymbol{e})^\top = 0.$$

また，$\boldsymbol{X}$ の第 1 列ベクトルが $\mathbf{1}_T = (1, \cdots, 1)^\top$ であるならば,

$$\mathbf{1}_T^\top \boldsymbol{e} = \sum_{t=1}^{T} e_t = 0.$$

以下, $\boldsymbol{X}$ の第 1 列ベクトルが $\mathbf{1}_T = (1, \cdots, 1)^\top$ であるとする．すなわち, (1.5) は,

$$y_t = \beta_1 + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + \epsilon_t, \quad t = 1, \cdots, T \tag{1.11}$$

とする．このとき，$\boldsymbol{y}$ の平均 $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ からの偏差の 2 乗和は,

$$\begin{aligned}(\boldsymbol{y} - \bar{y}\mathbf{1}_T)^\top(\boldsymbol{y} - \bar{y}\mathbf{1}_T) &= (\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T + \boldsymbol{e})^\top(\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T + \boldsymbol{e}) \\ &= (\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T)^\top(\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T) + 2\boldsymbol{e}^\top(\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T) + \boldsymbol{e}^\top\boldsymbol{e} \\ &= (\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T)^\top(\hat{\boldsymbol{y}} - \bar{y}\mathbf{1}_T) + \boldsymbol{e}^\top\boldsymbol{e}.\end{aligned}$$

すなわち,

$$\sum_{t=1}^{T}(y_t - \bar{y})^2 = \sum_{t=1}^{T}(\hat{y}_t - \bar{y})^2 + \sum_{t=1}^{T} e_t^2.$$

定義 1.5（決定係数）

$$R^2 = \frac{\sum_{t=1}^{T}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2} = 1 - \frac{\sum_{t=1}^{T} e_t^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

を**決定係数**という．

決定係数は，$k(=$ 説明変数の個数$)$ の単調増加であるため，これを補正した

$$\bar{R}^2 = 1 - \frac{\sum_{t=1}^{T} e_t^2/(T-k)}{\sum_{t=1}^{T}(y_t - \bar{y})^2/(T-1)}$$

を**修正済決定係数**という．

1.2.2　線形回帰式の検定

補題 1

$\boldsymbol{u} \sim N_k(\mathbf{0}, \boldsymbol{I}_k)$, A を $\mathrm{rank}(\boldsymbol{A}) = r \leq k$ の $k \times k$-実対称べキ等行列とすると,

$$\boldsymbol{u}^\top \boldsymbol{A}\boldsymbol{u} \sim \chi^2(r).$$

証明　$\boldsymbol{A}$ は、実対称行列であるから，λ_i, $i = 1, \cdots, k$ を $\boldsymbol{A}$ の固有値とすると,

$$\boldsymbol{\Lambda} := \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_k \end{pmatrix} = \boldsymbol{C}^\top A \boldsymbol{C}$$

となる直交行列 $\boldsymbol{C}$ が存在する (岩城 (2012) 定理 7.9 参照)．また, $\boldsymbol{A}$ が $\mathrm{rank}(\boldsymbol{A}) = r$

のベキ等行列であることから，一般性を失うこと無く，

$$\lambda_i = \begin{cases} 1 & i = 1, \cdots, r \\ 0 & i = r+1, \cdots, k \end{cases} \tag{1.12}$$

とできる．

∵ $\boldsymbol{A}$ の固有値と固有ベクトルを各々，λ と $\boldsymbol{x}$ とすると，各々の定義から，

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x},\ \boldsymbol{x} \neq \boldsymbol{0}.$$

A がベキ等とすると，

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\lambda\boldsymbol{x} = \lambda\boldsymbol{A}\boldsymbol{x} = \lambda^2\boldsymbol{x}. \\ \therefore \qquad \lambda\boldsymbol{x} &= \lambda^2\boldsymbol{x} \iff \lambda(\lambda-1)\boldsymbol{x} = \boldsymbol{0}. \end{aligned}$$

∴ λ は，0 か 1.

一方，$\boldsymbol{C}$ の各列は互いに直交するので $\mathrm{rank}(\boldsymbol{C}) = k$ であり，$\mathrm{rank}(\boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\top) = \mathrm{rank}(\boldsymbol{\Lambda})$. ∴ tr をトレース（= 対角成分の和）として，

$$r = \mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\top) = \mathrm{rank}(\boldsymbol{\Lambda}) = \mathrm{tr}(\boldsymbol{\Lambda}).$$

∴ (1.12) を得る．

いま，$\boldsymbol{z} = \boldsymbol{C}^\top\boldsymbol{u}$ とすると，期待値の線形性から，

$$\begin{aligned} \mathbb{E}[\boldsymbol{z}] &= \boldsymbol{C}^\top\mathbb{E}[\boldsymbol{u}] = \boldsymbol{0}, \\ \mathbb{E}[\boldsymbol{z}\boldsymbol{z}^\top] &= \boldsymbol{C}^\top\mathbb{E}[\boldsymbol{u}\boldsymbol{u}^\top]\boldsymbol{C} = \boldsymbol{C}^\top\boldsymbol{I}_k\boldsymbol{C} = \boldsymbol{I}_k. \end{aligned}$$

∴ $\boldsymbol{z} \sim N_k(\boldsymbol{0}, \boldsymbol{I}_k)$.

以上の議論から，

$$\begin{aligned} \boldsymbol{u}^\top\boldsymbol{A}\boldsymbol{u} &= \boldsymbol{z}^\top\boldsymbol{C}^\top\boldsymbol{A}\boldsymbol{C}\boldsymbol{z} \\ &= \boldsymbol{z}^\top\boldsymbol{C}^\top\boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\top\boldsymbol{C}\boldsymbol{z} = \boldsymbol{z}^\top\boldsymbol{\Lambda}\boldsymbol{z} = \sum_{t=1}^{r} z_i^2 \sim \chi^2(r). \end{aligned}$$

□

残差平方和 $SSE = \sum_{t=1}^{T} e_t^2$ については，次が成立する．

命題 1.1（残差平方和の分布）

$$\frac{SSE}{\sigma^2} \sim \chi^2(T-k).$$

証明

$$\begin{aligned} \boldsymbol{e} &= \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} \\ &= \boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} \end{aligned}$$

$$
\begin{aligned}
&= \left(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right)\boldsymbol{y} \\
&= \left(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right)(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\epsilon}) \\
&= \left(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right)\boldsymbol{\epsilon}. \qquad (1.13)
\end{aligned}
$$

容易に確かめられるように，$\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$ は対称ベキ等行列．かつ，

$$
\begin{aligned}
\operatorname{tr}\left(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right) &= \operatorname{tr}(\boldsymbol{I}_T) - \operatorname{tr}\left(\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right) \\
&= \operatorname{tr}(\boldsymbol{I}_T) - \operatorname{tr}\left((\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\right)^{5)} \\
&= \operatorname{tr}(\boldsymbol{I}_T) - \operatorname{tr}(\boldsymbol{I}_k) = T-k.
\end{aligned}
$$

$\therefore$補題 1 において，$\boldsymbol{A} = \boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$, $\boldsymbol{u} = \frac{\boldsymbol{\epsilon}}{\sigma}$ とすれば，題意を得る．
□

5) $\boldsymbol{A} = (a_{ij})$, $\boldsymbol{B} = (b_{ij})$ を各々 $T \times k$, $k \times T$ 行列とすると，

$$
\begin{aligned}
\operatorname{tr}(\boldsymbol{AB}) &= \sum_{i=1}^{T}\sum_{l=1}^{k} a_{il}b_{li} \\
&= \sum_{l=1}^{k}\sum_{i=1}^{T} b_{li}a_{il} \\
&= \operatorname{tr}(\boldsymbol{BA}).
\end{aligned}
$$

補題 2

$\hat{\boldsymbol{\beta}}$ と $\boldsymbol{e}$ は独立.

証明　$\hat{\boldsymbol{\beta}}$ と $\boldsymbol{e}$ が正規分布従い，無相関であることを示す．(1.13) とその導出と同様にして，

$$
\begin{aligned}
\hat{\boldsymbol{\beta}} &= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} = \boldsymbol{\beta} + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\epsilon}, \qquad (1.14) \\
\boldsymbol{e} &= (\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top)\boldsymbol{\epsilon}.
\end{aligned}
$$

どちらも $\boldsymbol{\epsilon} \sim N_T(\boldsymbol{0}, \sigma^2\boldsymbol{I}_T)$ の線形変換なので $(\hat{\boldsymbol{\beta}}, \boldsymbol{e})$ は正規分布に従う．

$$
\begin{aligned}
cov(\hat{\boldsymbol{\beta}}, \boldsymbol{e}) &= cov\left(\boldsymbol{\beta} + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\epsilon}, \left(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right)\boldsymbol{\epsilon}\right) \\
&= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top](\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top)^\top \\
&= \sigma^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top(\boldsymbol{I}_T - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top)^\top \\
&= \boldsymbol{0}.
\end{aligned}
$$

□

命題 1.2（回帰係数の検定統計量）

$\left(\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1}$ の (i,i) 要素を ξ_i とし，$s^2 = \frac{SSE}{T-k}$ とすると，

$$
\frac{\hat{\beta}_i - \beta_i}{s\sqrt{\xi_i}} \sim \mathrm{t}_{T-k}.
$$

証明　(1.14) より，$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \left(\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1}\boldsymbol{X}^\top\boldsymbol{\epsilon} \sim N_k\left(\boldsymbol{0}, \sigma^2\left(\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1}\right)$ であるから，$\frac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{\xi_i}} \sim N(0,1)$．あとは，補題 2 と命題 1.1 より，題意が成立する．□

1.2.3　回帰係数の F 検定

線形回帰式 (1.5) に対して，

$$\text{帰無仮説 } H_0 : \beta_{k-m+1} = \cdots = \beta_k = 0$$
$$\text{対立仮説 } H_1 : \text{それ以外}$$

の検定を考える.

定義 1.6（F 分布）

$X \sim \chi^2(m)$ と $Y \sim \chi^2(n)$ が互いに独立であるとき，$\frac{X/m}{Y/n}$ は，**自由度** (m,n) **の F 分布**にしたがうといい，このとことを，$\frac{X/m}{Y/n} \sim \mathrm{F}(m,n)$ で表す.

定理 1.7F 分布確率密度関数

自由度 (n_1, n_2) の t 分布の確率密度関数 $f(x; n_1, n_2)$ は次式で与えられる.

$$f(x, n_1, n_2) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \frac{x^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}, \quad x \in \mathbb{R}_+. \tag{1.15}$$

証明　$U \sim \chi^2(n_1)$，$V \sim \chi^2(n_2)$ として，U, V が互いに独立であるとすると，(U, V) の同時密度関数 $g(u, v)$ は，(1.1) より，

$$g(u, v) = \frac{1}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}} u^{\frac{n_1}{2}-1} v^{\frac{n_2}{2}-1} \mathrm{e}^{-\frac{u+v}{2}}, \quad (u, v) \in \mathbb{R}_+^2.$$

ここで，$x = \frac{u/n_1}{v/n_2}$，$y = v$ として，変数置換すると，ヤコビアンが $|J| = \frac{n_1}{n_2}y$ となるので，$(X, Y) = \left(\frac{U/n_1}{V/n_2}, V\right)$ の同時確率密度関数 $h(x, y)$ は,

$$\begin{aligned}
h(x, y) &= g\left(\frac{n_1}{n_2}xy, y\right)|J| \\
&= \frac{1}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}} \left(\frac{n_1}{n_2}xy\right)^{\frac{n_1}{2}-1} y^{\frac{n_2}{2}-1} \exp\left(-\frac{y}{2}\left(\frac{n_1}{n_2}x+1\right)\right), \\
&\quad (x, y) \in \mathbb{R}_+^2.
\end{aligned}$$

X の周辺確率密度関数は,

$$\begin{aligned}
f(x) &= \int_0^\infty h(x, y)\mathrm{d}y \\
&= \int_0^\infty \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}} y^{\frac{n_1+n_2}{2}-1} \exp\left(-\frac{y}{2}\left(\frac{n_1}{n_2}x+1\right)\right)\mathrm{d}y.
\end{aligned}$$

さらにここで，$z = \frac{y}{2}\left(\frac{n_1}{n_2}x+1\right)$ と変数置換すると,

$$f(x) = \int_0^\infty \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}} \left(\frac{2z}{\frac{n_1}{n_2}x+1}\right)^{\frac{n_1+n_2}{2}-1} \mathrm{e}^{-z} \left(\frac{2}{\frac{n_1}{n_2}x+1}\right)\mathrm{d}z$$

$$= \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\left(1+\frac{n_1 x}{n_2}\right)^{\frac{n_1+n_2}{2}}} \int_0^\infty z^{\frac{n_1+n_2}{2}-1} \mathrm{e}^{-z} \mathrm{d}z$$

$$= \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \frac{x^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1 x}{n_2}\right)^{\frac{n_1+n_2}{2}}}, \quad x \in \mathbb{R}_+.$$

∴題意を得る. □

帰無仮説が成立するとした場合の回帰式

$$y_t = \beta_1 x_{1t} + \cdots + \beta_{k-m} x_{(k-m)t} + \epsilon_t, \quad t = 1, \cdots, T$$

の残差平方和を SSE_0 とすると，次が成立する.

定義 1.3（F 統計量）

$$\frac{(SSE_0 - SSE)/m}{SSE/(T-k)} \sim \mathrm{F}(m, T-k)$$

証明　命題 1.1 より，$SSE_0 \sim \chi^2(T-(k-m))$ となることに注意すると，今までの議論と同様にして，題意を得る. □

例 1.1

$m = k-1$ の場合，すなわち，

$$\begin{aligned} &\text{帰無仮説 } H_0 : \beta_2 = \cdots = \beta_k = 0 \\ &\text{対立仮説 } H_1 : \text{それ以外} \end{aligned}$$

の場合.

H_0 が正しいとすると，残差は，$\boldsymbol{y} - \bar{y}\mathbf{1}_T$ となるので，決定係数は，

$$R^2 = 1 - \frac{SSE}{SSE_0}$$

となる. これより，

$$\frac{(SSE_0 - SSE)/(k-1)}{SSE/(T-k)} = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \sim \mathrm{F}(k-1, T-k).$$

参考文献

1) 岩城秀樹『Maximaで学ぶ経済・ファイナンス基礎数学』共立出版（2012）