目　次

2.　線形推測論 ······ 69
2.1　射影行列と逆行列 ······ 69
2.2　カイ 2 乗分布 ······ 74
2.3　フィッシャー・コクランの定理 ······ 77
2.4　t 分布と F 分布 ······ 82
2.5　ガウス・マルコフモデル ······ 84
2.6　仮説検定 ······ 92
2.7　平均の検定 ······ 96
2.8　重回帰分析 ······ 97
2.9　一元配置 ······ 104
2.10　二元配置 ······ 108

2

線形推測論

正規分布から派生する確率分布（カイ 2 乗分布，t 分布，F 分布）についてまとめる．ガウス・マルコフモデルの推定，検定統計量について述べ，線形モデルに応用する．

2.1　射影行列と逆行列

$\mathbf{R}^n$ が部分空間 U, V によって $\mathbf{R}^n = U \oplus V$ と直和分解されているとする．$\mathbf{R}^n$ の各元 x は $x = u + v$ $(u \in U, v \in V)$ と一意に分解されるが，$Px = u$ で定まる線形写像 $P : \mathbf{R}^n \to \mathbf{R}^n$ を V に沿った U への**射影** (射影子 projection) と呼ぶ．

P が射影であれば $P \circ P = P$ である．逆に $\mathbf{R}^n$ の線形写像 P がこの性質をもてば，$U = \{x; Px = x\}$ および $V = \{x; Px = 0\}$ とすれば，$\mathbf{R}^n = U \oplus V$ であり，P は V に沿った U への射影になる．

$\mathbf{R}^n$ の標準基底に関する P の行列表現も同じ記号 P で表すことにする．射影を表現する n 次行列を**射影行列**と呼ぶ．射影の性質 $P \circ P = P$ は行列のべき等条件 $P^2 = P$ と同値である．したがって，n 次正方行列 P が射影行列であるための必要十分条件は $P^2 = P$ となることである．I_n を n 次単位行列とする．

命題 1. P を n 次正方行列とするとき以下の条件は同値である：
(a) P は射影行列．
(b) P は $V = \{x; Px = 0\}$ に沿った $U = \{x; Px = x\}$ への射影に対応する

射影行列.
(c) P はべき等: $P^2 = P$.
(d) n 次正則行列 Q と r $(0 \leq r \leq n)$ が存在して

$$Q^{-1}PQ = \mathrm{diag}[\overbrace{1,...,1}^{r}, \overbrace{0,...,0}^{n-r}]$$

となる.
(e) $\mathrm{rank}\ P + \mathrm{rank}\ (I_n - P) = n$.
さらにこのとき $\mathrm{rank}\ P = \mathrm{tr}\ P = r$ である.

証明. $(b) \Rightarrow (a)$ は自明で, $(a) \Rightarrow (c) \Rightarrow (b)$ はすでに述べた. (b) を仮定する. U の基底 $x_1, ..., x_r$ と V の基底 $x_{r+1}, ..., x_n$ を取り, $Q = [x_1, ..., x_n]$ とすると, Q は正則であり, (d) が成り立つ. さらに, (d) ならば (e) は明らか. 最後に (e) を仮定する. $U_* = P\mathbf{R}^n, V_* = (I_n - P)\mathbf{R}^n$ とすると $\mathbf{R}^n = U_* + V_*$ であるから, $\dim\ (U_* \cap V_*) \equiv \dim\ U_* + \dim\ V_* - \dim\ (U_* + V_*) = n - n = 0$, すなわち $U_* \cap V_* = \{0\}$ となる. $P(I_n - P)\mathbf{R}^n = (I_n - P)P\mathbf{R}^n \subset U_* \cap V_* = \{0\}$ であるから (c) が示された. □

$\mathbf{R}^n$ に標準内積を考える：$x, y \in \mathbf{R}^n$ に対して $x \cdot y = x'y$. x と y が直交するとは $x \cdot y = 0$ となることで, これを $x \perp y$ と表す. $\mathbf{R}^n$ の部分空間 U に対して, $U^\perp = \{x \in \mathbf{R}^n;\ x \perp y\ (\forall y \in U)\}$ を U の**直交補空間**と呼ぶ. このとき $\mathbf{R}^n$ は $\mathbf{R}^n = U \oplus U^\perp$ と直和分解される. $U^\perp$ に沿った U への射影 P を U への**直交射影** (orthogonal projection) と呼び, 対応する行列を**直交射影行列**と呼ぶ.

命題 2. n 次正方行列 P が直交射影行列になるための必要十分条件は $P^2 = P$ かつ $P' = P$ となることである.

証明. P が直交射影行列であるとする. 射影行列であることから $P^2 = P$ は成り立つ. つぎに x_1, x_2 を $\mathbf{R}^n$ の任意の元とし, その分解を $x_i = u_i + v_i$

$(u_i \in U,\ v_i \in U^{\perp})$ とする．定義より $Px_i = u_i$ であるから，$x_1'P'x_2 = (Px_1)'x_2 = u_1'x_2 = u_1'(u_2+v_2) = u_1'u_2 = (u_1+v_1)'u_2 = x_1'Px_2$．したがって $P' = P$ である．

逆に $P^2 = P$ と $P' = P$ を仮定する．U, V を命題 1(b) のように取ると，$P^2 = P$ より $\mathbf{R}^n = U \oplus V$ であり，P は V に沿った U への射影に対応する行列であった．さらに，任意の $u \in U,\ v \in V$ に対して $u'v = (Pu)'v = u'P'v = u'Pv = 0$ であって，$V = U^{\perp}$ となるので，P は直交射影行列である．□

問 1. P を $\mathbf{R}^n$ の部分空間 U への直交射影とするとき，$y \in \mathbf{R}^n$ に対して，$|y - Py|^2 = y'(I_n - P)y = \min_{x \in U} |y - x|^2$ であることを示せ.

行列 A の列ベクトルで張られる線形空間を $L[A]$ と書く.

問 2. (i) $L[A] \subset L[B]$ であることと，ある行列 C が存在して $A = BC$ となることは同値であることを示せ．(ii) $L[A] = L[AA']$ となることを示せ．［ベクトル x が $x'AA' = 0$ を満たせば $x'AA'x = 0$，したがって $x'A = 0$ であるから $L[AA']^{\perp} \subset L[A]^{\perp}$.］

$L[A]$ への射影を表現するために逆行列の一般化をしておくと便利である.

定義 1. $m \times n$ 行列 A に対して $AA^-A = A$ を満足する $n \times m$ 行列 A^- を A の**一般 (化) 逆行列** (generalized inverse matrix) と呼ぶ.

A が正方行列で正則であれば A の一般逆行列は逆行列 A^{-1} に他ならない.

任意の $m \times n$ 行列 A に対して一般逆行列が存在する．実際，適当な $m \times m$ 正則行列 B と $n \times n$ 正則行列 C が存在して，

$$D = BAC = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \qquad (m \times n \text{ 行列})$$

とできるが，$A^- = CD'B$ とすればよい．

一般逆行列は一般に一意ではない．たとえば，

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

とするとき，A 自身も B も A の一般逆行列である．

X を $n \times p$ 行列とする．$\mathbf{R}^n$ から $L[X]$ への直交射影を P_X とし，標準基底に関する表現行列も P_X で表す．

定理 1. $X'X$ の任意の一般逆行列 $(X'X)^-$ に対して，$P_X = X(X'X)^-X'$.

証明．$Q = X(X'X)^-X'$ と置く．$L[X'] = L[X'X]$ であるから，ある行列 B に対して $X' = X'XB$ となる．このとき，

$$Q = B'X'X(X'X)^-X'XB = B'X'XB = B'X' \tag{2.1}$$

であるから，とくに $QX = B'X'X = X$ であり，$L[Q] = L[X]$．また，$Q' = B'X'XB = Q$ である．さらに，$Q^2 = QQ' = B'X'XB = Q$. つまり Q は $L[X]$ への直交射影であり，$Q = P_X$ である．□

$p \times q$ 行列 H が $L[H] \subset L[X']$，rank $H = q$ を満足すると仮定する．$L[X|H] = \{X\gamma; H'\gamma = 0\}$ とし，$L[X|H]$ への標準内積に関する直交射影を $P_{X|H}$ と表す．$r = \mathrm{rank}\ X$ とする．

定理 2. $(X'X)^-$ を（任意の）$X'X$ の一般逆行列とし，$G = (X'X)^-H$ とする．このとき，$G'X'XG$ は正則であり，$P_{X|H}$ は

$$P_{X|H} = P_X - XG(G'X'XG)^{-1}G'X'$$

と表示される．また，rank $P_{X|H} = r - q$ である．

証明．まず，$L[H] \subset L[X'] = L[X'X]$ よりある行列 B によって $H = X'XB$ と書け，$X'XG = H$ だから，rank $(G'X'XG) = $ rank $(XG) \geq$ rank $(X'XG) = $ rank $H = q$. したがって $G'X'XG$ は正則である．

$Q = P_X - XG(G'X'XG)^{-1}G'X'$ とし，$Q = P_{X|H}$ を示す．$\Gamma = (X'X)^{-}X' - G(G'X'XG)^{-1}G'X'$ とおくと，$Q = X\Gamma$． $H = X'XG$ と $P_X XG = XG$ によって，

$$H'\Gamma = G'X' - G'X'XG(G'X'XG)^{-1}G'X' = O$$

だから $L[Q] = L[X\Gamma] \subset L[X|H]$. 逆に $H'\gamma = 0$ なる任意の γ に対して，

$$\begin{aligned} X\gamma &= P_X X\gamma \\ &= P_X X\gamma - XG(G'X'XG)^{-1}G'X'X\gamma \qquad (G'X'X\gamma = H'\gamma = 0) \\ &= QX\gamma \end{aligned}$$

だから $L[X|H] \subset L[Q]$，したがって，$L[Q] = L[X|H]$ である．さらに，Q が対称べき等であることも容易にわかるので $Q = P_{X|H}$ である．また，$\mathrm{rank}\, Q = \mathrm{tr}\, Q = \mathrm{tr}\, P_X - \mathrm{tr}\,(XG(G'X'XG)^{-1}G'X') = \mathrm{tr}\, P_X - \mathrm{tr}\,(G'X'XG(G'X'XG)^{-1}) = r - q$ である．□

ブロック化された行列の逆行列の表現に触れておく．A を $p \times p$, B を $p \times q$, C を $q \times p$, D を $q \times q$ 行列とする．$(p+q) \times (p+q)$ 行列 M を

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

とする．さらに，D が正則であるとき $F = A - BD^{-1}C$，A が正則であるとき $G = D - CA^{-1}B$ とおく．

命題 3. (a) D が正則であるとき，$|M| = |F||D|$. とくに，D, M が正則であることと D, F が正則であることは同値である．
(b) D と M が共に正則であるとき，

$$M^{-1} = \begin{bmatrix} F^{-1} & -F^{-1}BD^{-1} \\ -D^{-1}CF^{-1} & D^{-1} + D^{-1}CF^{-1}BD^{-1} \end{bmatrix}.$$

(c) A が正則であるとき，$|M| = |A||G|$．とくに，A, M が正則であることと A, G が正則であることは同値である．
(d) A と M が共に正則であるとき，

$$M^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BG^{-1}CA^{-1} & -A^{-1}BG^{-1} \\ -G^{-1}CA^{-1} & G^{-1} \end{bmatrix}.$$

証明．D が正則であるとき，等式

$$\begin{bmatrix} I_p & -BD^{-1} \\ O & I_q \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_p & O \\ -D^{-1}C & I_q \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & O \\ O & D \end{bmatrix}$$

が成り立つ．両辺の行列式を取って (a) を得る．さらに M も正則であれば，(a) から $|F| \neq 0$，したがって F は正則であり，上の等式から (b) を得る．(c),(d) も同様の議論によるか，あるいは行と列の順序を変える変換を用いて (a), (b) に帰着させることによって，示すことができる．□

問 3. $N = \begin{bmatrix} A & B \\ B' & D \end{bmatrix}$ を対称行列とする．$F = A - BD^{-}B'$，$G = D - B'A^{-}B$ とおく．$L[B'] \subset L[D]$ のとき

$$\begin{bmatrix} F^{-} & -F^{-}BD^{-} \\ -D^{-}B'F^{-} & D^{-} + D^{-}B'F^{-}BD^{-} \end{bmatrix}$$

は N の一般逆行列であることを示せ．また，$L[B] \subset L[A]$ のとき

$$\begin{bmatrix} A^{-} + A^{-}BG^{-}B'A^{-} & -A^{-}BG^{-} \\ -G^{-}B'A^{-} & G^{-} \end{bmatrix}$$

は N の一般逆行列であることを示せ．

2.2 カイ２乗分布

正規分布の２次形式の分布を考えよう．ガンマ分布 $G(\alpha, \nu)$ の確率密度関数 (**??**) を $g(x; \alpha, \nu)$ で表す．

命題 4. X_j $(j=1,...,k)$ を実確率変数とする．X_1 が正規分布 $N(\mu,1)$，X_j $(j=2,...,k)$ が正規分布 $N(0,1)$ に従い，$X_1,...,X_k$ は独立であるとする．このとき，$Y=\sum_{j=1}^k X_j^2$ の確率密度関数は

$$f(x)=1_{\{x>0\}}e^{-\frac{\mu^2}{2}}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\mu^2}{2}\right)^r g(x;\frac{1}{2},r+\frac{k}{2}),$$

特性関数は

$$\begin{aligned}\varphi(u)&=e^{-\frac{\mu^2}{2}}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\mu^2}{2}\right)^r(1-2iu)^{-r-\frac{k}{2}}\\&=(1-2iu)^{-\frac{k}{2}}\exp\left(\frac{\mu^2 iu}{1-2iu}\right)\end{aligned}$$

で与えられる．

証明．$Z=X_1^2$ の確率密度関数を計算しよう．まず分布関数は，$z>0$ に対して，

$$\begin{aligned}P[Z\le z]&=P[|X_1|\le\sqrt{z}]\\&=\int_{-\sqrt{z}}^0\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\mu)^2}dx+\int_0^{\sqrt{z}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\mu)^2}dx\\&=\int_0^z\frac{1}{2\sqrt{2\pi s}}e^{-\frac{1}{2}(\sqrt{s}+\mu)^2}ds+\int_0^z\frac{1}{2\sqrt{2\pi s}}e^{-\frac{1}{2}(\sqrt{s}-\mu)^2}ds.\end{aligned}$$

したがって，Z の確率密度関数 f_Z は

$$\begin{aligned}f_Z(z)&=\frac{1}{2\sqrt{2\pi z}}e^{-\frac{1}{2}(z+\mu^2)}\left[e^{-\mu\sqrt{z}}+e^{\mu\sqrt{z}}\right]\\&=e^{-\frac{1}{2}\mu^2}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\mu^2}{2}\right)^r g(z;\frac{1}{2},r+\frac{1}{2})\end{aligned}$$

となる．

$X_2^2,...,X_k^2$ は独立にガンマ分布 $G(1/2,1/2)$ に従うから，再生性 (第**??**節) によって，$\sum_{j=2}^k X_j^2\sim G(1/2,(k-1)/2)=\chi^2(k-1)$ である．$Y=Z+\sum_{j=2}^k X_j^2$

の確率密度関数 f は，上の式と $g(\cdot;\frac{1}{2},\frac{k-1}{2})$ とのたたみこみを取って，$y>0$ に対して，

$$f(y)=e^{-\frac{1}{2}\mu^2}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\mu^2}{2}\right)^r g(y;\frac{1}{2},r+\frac{k}{2})$$

であることがわかり前半が示された．この密度関数の特性関数は

$$\varphi(u)=e^{-\frac{1}{2}\mu^2}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\mu^2}{2}\right)^r(1-2iu)^{-(r+\frac{k}{2})}$$

であり，簡単な計算によって後半を得る．□

定義 2. 命題 4 の $\sum_{j=1}^{k}X_j^2$ の分布を自由度 k，非心率（非心度）$\delta=\mu^2$ の**非心カイ２乗分布** (noncentral chi-square distribution with k degree of freedom and noncentrality parameter δ) と呼び，$\chi^2(k,\delta)$ で表す．

$\delta=0$ のとき，非心カイ 2 乗分布はカイ 2 乗分布になる：$\chi^2(k,0)=\chi^2(k)$. 便宜上，自由度 0 の非心カイ 2 乗分布は $x=0$ 上の 1 点分布とし，そのとき非心率は $\delta=0$ とする．

命題 4 より，$k>0$ のとき，非心カイ 2 乗分布 $\chi^2(k,\delta)$ の確率密度関数は

$$f(y)=e^{-\frac{1}{2}\delta}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\delta}{2}\right)^r g(y;\frac{1}{2},r+\frac{k}{2})\qquad(y>0),$$

特性関数は

$$\varphi(u)=(1-2iu)^{-\frac{k}{2}}\exp\left(\frac{\delta iu}{1-2iu}\right)\tag{2.2}$$

となる．

次の二つの結果は特性関数 (2.2) の形から容易に得られる．

定理 3. $X_j\sim N(\mu_j,1)$ $(j=1,...,k)$, $X_1,...,X_k$ は独立とする．このとき，$\sum_{j=1}^{k}X_j^2\sim\chi^2(k,\delta)$ となる．ここで，$\delta=\sum_{j=1}^{k}\mu_j^2$.

定理 4. 確率変数 X_j $(j = 1, ..., k)$ が独立に $\chi^2(k_j, \delta_j)$ に従うとき，$\sum_{j=1}^k X_j$ は $\chi^2(\sum_{j=1}^k k_j, \sum_{j=1}^k \delta_j)$ に従う．

2.3　フィッシャー・コクランの定理

定理 5. A_i $(i = 1, ..., k)$ を n 次正方行列とし，

$$A_1 + \cdots + A_k = I_n \tag{2.3}$$

を満足すると仮定する．このとき，次の $(a), (b), (c), (d)$ は同値である：

(a) A_i はべき等である：$A_i^2 = A_i$ $(i = 1, ..., k)$.
(b) $A_i A_j = O$ $(i \neq j)$.
(c) $V_i = A_i\ \mathbf{R}^n$ とするとき $\mathbf{R}^n$ が $V_1 \oplus \cdots \oplus V_k = \mathbf{R}^n$ と直和分解される．
(d) $\sum_{i=1}^k \mathrm{rank} A_i = n$.

証明．$(b) \Rightarrow (a)$: (2.3) の両辺に右から A_i をかければよい．
$(a) \Rightarrow (d)$: A_i はべき等だから rank A_i = tr A_i. (2.3) の両辺のトレースを取って (d) を得る．
$(d) \Rightarrow (c)$: $V_1 + \cdots + V_k = \mathbf{R}^n$ となることは (2.3) からわかる．直和になることは，条件 (d) によって $\sum_{i=1}^k \dim V_i = n$ となることからいえる．*1)
$(c) \Rightarrow (b)$: (2.3) の両辺の右辺から A_j をかけると任意の $x \in \mathbf{R}^n$ に対して，

$$A_1 A_j x + \cdots + A_j^2 x + \cdots + A_k A_j x = A_j x$$

となるが，左辺の $A_i A_j x$ は V_i の元で，直和分解であることから，$A_i A_j x = 0$ $(i \neq j)$ でなければならない．したがって (b) を得た．□

$\mu = [\mu_1, ..., \mu_n]' \in \mathbf{R}^n$, $Y = [Y_1, ..., Y_n]' \sim N_n(\mu, I_n)$ とする．

*1)　たとえば佐武 [7]，p.98 を参照せよ．

定理 6. (フィッシャー・コクランの定理) *2) A_i $(i=1,...,k)$ を n 次対称行列で $A_1+\cdots+A_k=I_n$ を満たすものとする．$Q_i=Y'A_iY$, rank $A_i=n_i$ $(i=1,...,k)$ とする．このとき次の (a)-(e) は同値である：

(a) ある $\delta_i \geq 0$ が存在して $Q_i \sim \chi^2(n_i,\delta_i)$ かつ $Q_1,...,Q_k$ は独立.

(b) $\sum_{i=1}^k n_i = n$.

(c) $A_i^2=A_i$ $(i=1,...,k)$.

(d) $A_iA_j=O$ $(i\neq j)$.

(e) $V_i=A_i\,\mathbf{R}^n$ とするとき $\mathbf{R}^n$ が $V_1\oplus\cdots\oplus V_k=\mathbf{R}^n$ と直和分解される.

さらにこのとき，$\delta_i=\mu'A_i\mu$，とくに $\sum_{j=1}^n \mu_j^2=\sum_{i=1}^k \delta_i$.

証明．$(b),(c),(d),(e)$ の同値性は定理 5 で示した．(b) を仮定する．$A_i^2=A_i$, $A_iA_j=O$ $(i\neq j)$ である．A_i は対称でべき等だからある n 次直交行列 P_i が存在して，

$$P_iA_iP_i'=D_i=\mathrm{diag}[\overbrace{1,...,1}^{n_i},\overbrace{0,...,0}^{n-n_i}]$$

とできる．$n_i=0$ のときは $A_i=O$ だから，$Y'A_iY\sim\delta_0$ （１点分布）となる．$n_i>0$ とする．$X_i=P_iA_iY$ と置くと，

$$\mathrm{Cov}[X_i,X_j]=P_iA_iA_j'P_j'=P_iA_iA_jP_j'=\delta_{ij}D_i,$$

さらに，

$$E[X_i]=P_iA_i\mu=D_iP_i\mu=\begin{bmatrix}\hat{D}_iP_i\mu\\0\end{bmatrix}$$

となる．ここで $\hat{D}_i$ は $n_i\times n$ 行列で

$$\hat{D}_i=[I_{n_i},\overbrace{O}^{n-n_i}]$$

*2) R.A.Fisher, W.G. Cochran

である．したがって定理??によって $X_1, ..., X_k$ は独立であり，

$$\hat{X}_i := \hat{D}_i X_i \sim N_{n_i}(\hat{D}_i P_i \mu, I_{n_i}), \qquad X_i = \begin{bmatrix} \hat{X}_i \\ 0 \end{bmatrix},$$

したがって，

$$Y' A_i Y = Y' A_i' P_i' P_i A_i Y = X_i' X_i = \hat{X}_i' \hat{X}_i \sim \chi^2(n_i, \delta_i)$$

であることがわかった．ここで $\delta_i = (\hat{D}_i P_i \mu)' \hat{D}_i P_i \mu = \mu' P_i' D_i P_i \mu = \mu' A_i \mu$ である．$X_i' X_i$ は独立であるから (a) が示された．

逆に (a) を仮定し (b) を示そう．$n_i = 0$ ならば $A_i = O$ であるから，このためには $n_i > 0$ と仮定しても一般性を失わない．$Y'Y = \sum_{i=1}^k Q_i$ と定理 4 より $Y'Y \sim \chi^2(\sum_{i=1}^k n_i, \sum_{i=1}^k \delta_i)$ である．いっぽう $Y'Y = \sum_{j=1}^n Y_j^2$ より $Y'Y \sim \chi^2(n, \mu'\mu)$．2 つのカイ 2 乗分布の特性関数が等しいから

$$\begin{aligned} &\left(1 - 2\sqrt{-1}u\right)^{-\frac{1}{2}\sum_{i=1}^k n_i} \exp\left(\frac{\sum_{i=1}^k \delta_i \sqrt{-1}u}{1 - 2\sqrt{-1}u}\right) \\ &= \left(1 - 2\sqrt{-1}u\right)^{-\frac{1}{2}n} \exp\left(\frac{\mu'\mu\sqrt{-1}u}{1 - 2\sqrt{-1}u}\right) \end{aligned}$$

となるが，$u \in \mathbf{R}$, $u \to \infty$ とすると，指数関数の分数部分は有界だから $\sum_{i=1}^k n_i = n$ でなければならず (b) が得られた．□

定理 7. A を n 次対称行列とする．ある $\delta \geq 0$, $k \in \mathbf{Z}_+$ に対して $Q = Y'AY$ が $\chi^2(k, \delta)$ に従うための必要十分条件は $A^2 = A$ となることである．さらにこのとき，$k = \mathrm{rank} A$, $\delta = \mu' A \mu$ である．

証明．$A^2 = A$ と仮定する．$(I_n - A)^2 = I_n - A$ となるから，分解 $A + (I_n - A) = I_n$ に定理 6 を適用して $Q \sim \chi^2(k, \delta)$, $k = \mathrm{rank} A$, $\delta = \mu' A \mu$ となることがわかる．

逆に $Q \sim \chi^2(k, \delta)$ とする．Y の分布のサポートが $\mathbf{R}^n$ 全体であるから，とくに A は非負値である．また $k = 0$ のときは $A = O$ であるか

ら $A^2 = A$ となるので，$k > 0$ の場合を考える．A に対して直交行列 P が存在して，$PAP' = \mathrm{diag}[\lambda_1, ..., \lambda_p, 0, ..., 0]$ となる．ここで，$\lambda_1, ..., \lambda_p$ $(p \le n)$ は A の正の固有値である．$X = [X_1, ..., X_n]' := PY$ とするとき，$Q = Y'AY = \sum_{i=1}^p \lambda_i X_i^2$ と表される．$X \sim N_n(P\mu, I_n)$ だから $\sum_{i=1}^p \lambda_i X_i^2$ は独立な p 個の非心カイ 2 乗分布の重み付き和になっていて，その特性関数は，ある $\delta_i \ge 0$ に対して，

$$\prod_{i=1}^p \left\{ \left(1 - 2\sqrt{-1}\lambda_i u\right)^{-\frac{1}{2}} \exp\left(\frac{\delta_i \sqrt{-1}\lambda_i u}{1 - 2\sqrt{-1}\lambda_i u} \right) \right\} \tag{2.4}$$

となる．

いっぽう，$Q \sim \chi^2(k, \delta)$ だったからその特性関数は

$$\left(1 - 2\sqrt{-1}u\right)^{-\frac{k}{2}} \exp\left(\frac{\delta\sqrt{-1}u}{1 - 2\sqrt{-1}u} \right) \tag{2.5}$$

であり，(2.4) と (2.5) が一致することになる．両者から導かれるキュムラントを比較する．まず，(2.4) の対数の $u = 0$ の近傍での展開は

$$\begin{aligned} &-\frac{1}{2}\sum_{i=1}^p \log\left(1 - 2\sqrt{-1}\lambda_i u\right) + \sum_{i=1}^p \frac{\delta_i\sqrt{-1}\lambda_i u}{1 - 2\sqrt{-1}\lambda_i u} \\ &= \frac{1}{2}\sum_{i=1}^p \sum_{r=1}^\infty \frac{(2\lambda_i)^r}{r}(\sqrt{-1}u)^r + \sum_{i=1}^p \sum_{r=0}^\infty \delta_i \lambda_i (2\lambda_i)^r (\sqrt{-1}u)^{r+1} \\ &= \sum_{r=1}^\infty \frac{(\sqrt{-1}u)^r}{r!} \sum_{i=1}^p \left[(r-1)! 2^{r-1} \lambda_i^r + r! \delta_i 2^{r-1} \lambda_i^r \right] \end{aligned}$$

となり，r 次キュムラントは $\sum_{i=1}^p \left[(r-1)!2^{r-1}\lambda_i^r + r!\delta_i 2^{r-1}\lambda_i^r\right]$ である．(2.5) に対しても同様にして，対応する r 次キュムラントを求めると $k(r-1)!2^{r-1} + r!\delta 2^{r-1}$ となり，結局，

$$\sum_{i=1}^p \lambda_i^r + r\sum_{i=1}^p \delta_i \lambda_i^r = k + r\delta \qquad (r \in \mathbf{N}) \tag{2.6}$$

を得る．λ_i の中で 1 つでも 1 より大きいものがあると (2.6) の左辺は $r \to \infty$ のとき指数関数的となり矛盾．したがって $\lambda_1, ..., \lambda_q = 1$, $\lambda_{q+1}, ..., \lambda_p < 1$ と

してよいが，(2.6) を r で割って $r \to \infty$ として $\sum_{i=1}^{q} \delta_i = \delta$ を得る．これを使って, (2.6) で $r \to \infty$ とすると $q = k$ がわかり，再び (2.6) に戻すと，

$$\sum_{i=q+1}^{p} \lambda_i^r + r \sum_{i=q+1}^{p} \delta_i \lambda_i^r = 0$$

となる．したがって，$p = q$ でなければならず，λ_i はすべて 1 であることが示された．このとき，$A = P' \mathrm{diag}[1, ..., 1, 0, ..., 0] P$ であって A はべき等である．□

定理 8. A_i $(i = 1, 2)$ を n 次対称行列とする．$Q_i = Y' A_i Y$ が非心カイ２乗分布に従うとする．このとき，Q_1 と Q_2 が独立であるための必要十分条件は $A_1 A_2 = O$ となることである．

証明．Q_1 と Q_2 が独立であるとき，$Q_1 + Q_2 = Y'(A_1 + A_2)Y$ も非心カイ２乗分布に従う（定理 4)．したがって，定理 7 によって，$A_1, A_2, A_1 + A_2$ がべき等であり，$A_1 A_2 + A_2 A_1 = O$ を得るが，A_1 を右からと左からと掛けたものを比較して $A_1 A_2 = A_2 A_1 = O$ となることがいえる．

逆に $A_1 A_2 = O$ であるとする．単位行列の分解

$$A_1 + A_2 + (I_n - A_1 - A_2) = I_n$$

において，右辺の各項は条件からべき等である．したがって定理 6 によって Q_1, Q_2 は独立である．□

問 4. A を n 次対称行列，Σ を n 次正値対称行列とし，Z が $N_n(\mu, \Sigma)$ に従うとする．このとき，$Z' A Z$ がある $k \in \mathbf{N}$ と $\delta \geq 0$ に対して $\chi^2(k, \delta)$ に従うための必要十分条件はなにか．[$A \Sigma A = A$．このとき $k = \mathrm{rank}\ A$, $\delta = \mu' A \mu$ である．]

補題 1. A, B を n 次対称行列とする．$A^2 = A$, $B^2 = B$, $A - B \geq O$ すなわち $A - B$ が非負値行列であるとする．このとき，$AB = BA = B$, $(A - B)^2 = A - B$ である．

証明．$A-B \geq O$, $B^2=B$ だから，$BAB \geq BBB = B$. ゆえに，

$$\mathrm{tr}(BAB) \geq \mathrm{tr}B = \mathrm{tr}(BB) = \mathrm{tr}(BAB) + \mathrm{tr}(B(I_n - A)B).$$

したがって，

$$0 \geq \mathrm{tr}(B(I_n - A)B) = \mathrm{tr}(((I_n - A)B)'(I_n - A)B) \geq 0$$

だから $(I_n - A)B = O$ である．転置して，$AB = BA = B$ を得る．$(A-B)^2 = A-B$ は容易に示される．□

定理 9. A_i $(i=0,1,2)$ を n 次対称行列，$A_0 = A_1 + A_2$ とし，$Q_i = Y'A_iY$ とする．$Q_0 \sim \chi^2(n_0, \delta_0)$, $Q_1 \sim \chi^2(n_1, \delta_1)$, $Q_2 \geq 0$ a.s. ならば $n_2 = n_0 - n_1$, $\delta_2 = \delta_0 - \delta_1$ に対して $Q_2 \sim \chi^2(n_2, \delta_2)$ であり，さらに Q_1 と Q_2 は独立である．

証明．定理 7 より，$A_i^2 = A_i$, $\mathrm{rank}A_i = n_i$, $\delta_i = \mu'A_i\mu$ $(i=0,1)$ である．$Q_2 \geq 0$ より $A_0 - A_1 \geq O$ だから，補題 1 から $A_2 = A_0 - A_1$ はべき等である．再び定理 7 より $Q_2 = Y'A_2Y \sim \chi^2(n_2, \delta_2)$, $n_2 = \mathrm{rank}A_2 = \mathrm{tr}A_2 = n_0 - n_1$, $\delta_2 = \mu'A_2\mu = \mu'(A_0 - A_1)\mu = \delta_0 - \delta_1$ となる．さらに，補題 1 から $A_1A_2 = A_1A_0 - A_1^2 = A_1 - A_1 = O$ だから，定理 8 から Q_1 と Q_2 は独立である．□

2.4 t 分布と F 分布

この節では正規分布から派生する 2 つの分布を定義する．

定義 3. Y が自由度 n のカイ 2 乗分布 $\chi^2(n)$ に従い，Z が正規分布 $N(\delta, 1)$ に従い，Y, Z が独立であるとする．このとき

$$X = \frac{Z}{\sqrt{Y/n}}$$

の分布を自由度 n，非心度 δ の**非心 t 分布** (noncentral t-distribution) と呼び $t(n,\delta)$ で表す．とくに $\delta=0$ のとき自由度 n の t **分布**と呼び，$t(n)$ で表す．

例 1. $X_1,...,X_n$ を正規分布 $N(\mu,\sigma^2)$ に従う独立確率変数列とする．U を第１行が $[1/\sqrt{n},...,1/\sqrt{n}]$ であるような n 次直交行列とする．これにはたとえば第 i 行を

$$\Big[\underbrace{\frac{1}{\sqrt{(i-1)i}},...,\frac{1}{\sqrt{(i-1)i}}}_{i-1},\ -\frac{i-1}{\sqrt{(i-1)i}},\ \underbrace{0,...,0}_{n-i}\Big] \qquad (i=2,...,n)$$

と取ればよい．$X=[X_1,...,X_n]'$ を $Y=[Y_1,...,Y_n]'=UX$ と変換すると，命題**??**によって，$Y_1=\sqrt{n}\bar{X}\sim N(\sqrt{n}\mu,\sigma^2)$, $Y_i\sim N(0,\sigma^2)$ $(i=2,...,n)$, さらに $Y_1,Y_2,...,Y_n$ は独立になる．S^2 を標本分散 $S^2=n^{-1}\sum_{j=1}^n(X_j-\bar{X})^2=n^{-1}\sum_{j=1}^n X_j^2-\bar{X}^2$ とすると，直交変換の性質によって $\sum_{j=1}^n X_j^2=\sum_{i=1}^n Y_i^2$ であるから，$S^2=n^{-1}\sum_{i=2}^n Y_i^2$ となる．したがって，nS^2/σ^2 は Y_1 と独立で自由度 $n-1$ の χ^2 分布に従い，

$$\frac{\sqrt{n-1}(\bar{X}-\mu)}{S}=\frac{Y_1-\sqrt{n}\mu}{\sigma}\Big/\sqrt{\frac{nS^2}{\sigma^2(n-1)}}$$

は自由度 $n-1$ の t 分布に従う．

Y,Z を定義 3 で与えられたものとする．変換 $(z,y)=T(x_1,x_2)=(x_1\sqrt{x_2}/\sqrt{n},x_2)$ を考えると，$(X_1,X_2)=T^{-1}(Z,Y)$ の同時密度関数は，$(x_1,x_2)\in\mathbf{R}\times(0,\infty)$ に対して，

$$p^*(x_1,x_2)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x_1\sqrt{x_2}}{\sqrt{n}}-\delta\right)^2\right\}\cdot\frac{x_2^{\frac{n}{2}-\frac{1}{2}}}{2^{n/2}\sqrt{n}\Gamma(n/2)}e^{-\frac{1}{2}x_2}.$$

X_1 の周辺確率密度関数を計算して，非心 t 分布の確率密度関数を得る：$x\in\mathbf{R}$ に対して，

$$p_{n,\delta}(x)=\frac{1}{\sqrt{\pi n}}e^{-\frac{\delta^2}{2}}\sum_{r=0}^{\infty}\frac{2^{r/2}}{r!}\frac{\Gamma((n+r+1)/2)}{\Gamma(n/2)}\left(\frac{\delta x}{\sqrt{n}}\right)^r\left(1+\frac{x^2}{n}\right)^{-(n+r+1)/2}.$$

とくに $\delta=0$ のとき $r=0$ の項のみ残り，t 分布の確率密度関数を表す．

定義 4. Y_1 が $\chi^2(m,\delta)$ に，Y_2 が $\chi^2(n)$ に従い，Y_1, Y_2 が独立であるとする．このとき，

$$X = \frac{Y_1/m}{Y_2/n}$$

の分布を自由度 m, n，非心度 δ の**非心 F 分布** (noncentral F-distribution) と呼び，$F(m,n,\delta)$ と表す．とくに $\delta = 0$ のとき自由度 m, n の F **分布**と呼び，$F(m,n)$ で表す．

非心 F 分布の確率密度関数は，$x \in (0,\infty)$ に対して，

$$p_{m,n,\delta}(x) = \sum_{r=0}^{\infty} e^{-\frac{1}{2}\delta} \left(\frac{\delta}{2}\right)^r \frac{(m/n)^{m/2+r}}{r!B(m/2+r,n/2)} \frac{x^{m/2+r-1}}{(1+mx/n)^{(m+n)/2+r}}$$

となる．とくに $\delta = 0$ のとき $r = 0$ の項のみ残り，それが F 分布の確率密度関数 $p_{m,n,0}$ になる．また，合流型超幾何関数（Kummer の関数）*1)

$$F(\alpha,\gamma;z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)}{\Gamma(\gamma+n)} \frac{z^n}{n!}$$

を用いて，

$$p_{m,n,\delta}(x) = e^{-\frac{1}{2}\delta} F\left(\frac{m+n}{2}, \frac{m}{2}; \frac{\delta}{2}\frac{mx/n}{1+mx/n}\right) p_{m,n,0}(x)$$

と表される．

2.5 ガウス・マルコフモデル

n 個の実確率変数 Y_i $(i=1,...,n)$ が

$$Y_i = x_i'\beta + \epsilon_i \qquad (i=1,...,n) \tag{2.7}$$

なる構造を持つとする．ここで，x_i，β は非確率的 p 次元列ベクトル，ϵ_i は平均零の等分散かつ無相関な確率変数：

$$E[\epsilon_i] = 0, \qquad E[\epsilon_i\epsilon_j] = \sigma^2\delta_{ij}$$

*1) たとえば犬井 [7] を参照.

とする．統計推測においては Y_i が i 番目の観測を表し，x_i は $E[Y_i]$ を説明する p 個の要因を並べた変数，$\beta \in \mathbf{R}^p$ は x_i の p 個の要因への $E[Y_i]$ の依存の程度を表すパラメータ，σ^2 は観測の分散を表すパラメータである．

n 次元確率変数 Y, ϵ, $n \times p$ 定行列 X を

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}, \quad X = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

と定める．このとき，(2.7) は次のように行列表示できる：

$$Y = X\beta + \epsilon, \quad E[\epsilon] = 0, \quad \mathrm{Var}[\epsilon] = \sigma^2 I_n. \tag{2.8}$$

モデル (2.8) を**ガウス・マルコフモデル**と呼ぶ．実際は実験のデザイン X は既知であるが，β, σ^2 は未知のパラメータであり，デザイン X に対応する観測 Y に基づいて $E[Y]$ と X の関係を規定するパラメータ β とデータのばらつきの指標 σ^2 について推測するのである．$E[Y]$ の β に関する線形性はもちろんここでの前提であるが，変量間の関係の表現の第 1 近似と思えばよい．また，デザイン X は根元的な要因の非線形関数とすることが可能なので，この前提は応用上は強い制約ではない．

未知パラメータ β の推定において，最小 2 乗法は最も基本的な方法である．

$$Q(\beta) = (Y - X\beta)'(Y - X\beta) = \sum_{i=1}^{n} (Y_i - x_i'\beta)^2 \tag{2.9}$$

とき，関数 Q を $\beta \in \mathbf{R}^p$ を動かして最小にする．Q を β に関して微分して 0 とおくことによって

$$X'X\beta = X'Y \tag{2.10}$$

を得る．β の方程式 (2.10) は**正規方程式** (normal equation) と呼ばれる．

$L[X'X] = L[X']$ であるから，正規方程式 (2.10) は常に解 $\hat{\beta}$ を持つ．いっぽう，解 $\hat{\beta}$ は $X'X$ が正則でなければ一意には定まらない．

定理 10. $Y \in \mathbf{R}^n$ とする．正規方程式 (2.10) の任意の解 $\hat{\beta}$ に対して，$X\hat{\beta} = P_X Y$．また，$|Y - X\hat{\beta}| = \min_{\beta \in \mathbf{R}^p} |Y - X\beta|$.

証明．定理 1，正規方程式と P_X の性質を用いて，$P_X Y = X(X'X)^- X'Y = X(X'X)^- X'X\hat{\beta} = P_X X\hat{\beta} = X\hat{\beta}$ となる．後半は，$(Y - X\hat{\beta})'X = Y'(I_n - P_X)X = 0$ から従う分解 $|Y - X\beta|^2 = |Y - X\hat{\beta}|^2 + |X(\hat{\beta} - \beta)|^2$ を用いて証明できる．(問 1 参照．)□

未知パラメータ β, σ^2 の関数 $g(\beta, \sigma^2)$ の推定を考える．推定量（パラメータの関数の推定に用いる観測の可測関数）はどの (β, σ^2) が真の値であっても統計的に「良い」推定値を与えるものでなければならない．次に述べる**不偏性** (unbiasedness) は推定量の望ましい性質の一つである．

特定の β と σ^2 によって Y の構造が (2.8) で与えられるときの，統計量 $\delta(Y)$ の期待値を $E_{\beta,\sigma^2}[\delta(Y)]$ と表す．$\delta(Y)$ がパラメータ (β, σ^2) の関数 $g(\beta, \sigma^2)$ の**不偏推定量**であるとは，任意の真値 (β, σ^2) に対して，$\delta(Y)$ の平均が $g(\beta, \sigma^2)$ となることである．すなわち，

$$E_{\beta,\sigma^2}[\delta(Y)] = g(\beta, \sigma^2) \qquad (\forall(\beta, \sigma^2) \in \Pi) \tag{2.11}$$

となることである．ここで Π は (β, σ^2) の動く範囲で，我々は多くの場合 $\Pi = \mathbf{R}^p \times (0, \infty)$ とする．以後 E_{β,σ^2} を単に E と略すことも多い．

我々は Y の構造として，平均と共分散しか規定していなかった．平均 $X\beta$，分散共分散行列 $\sigma^2 I_n$ の分布は無限にあり，そのような分布に関する平均を E_{β,σ^2} と書いても，これは多義であって，たとえば $E_{\beta,\sigma^2}[Y_1^3]$ の値を定めることはできない．この章で扱う不偏推定量 $\delta(Y)$ は Y の 2 次以下の多項式で構成されるので，そのような不定性は生じない．

記号 E_{β,σ^2} は便利であるが，どの確率測度に関する期待値か明確でなく，(2.11) の左辺の意味が曖昧である．より明瞭で一般的な不偏推定量の定義を述べよう．[*1)]

---

*1) この章の理解のためには，次の定義とそのすぐ後の議論は飛ばしてもよい．

定義 5. 可測空間 $(\mathcal{X}, \mathcal{B})$ 上の確率測度の族 $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$ と写像 $g : \Theta \to \mathbf{R}^k$ が与えられているとする．$g(\theta)$ の推定量 $\delta : \mathcal{X} \to \mathbf{R}^k$ が

$$\int_{\mathcal{X}} \delta(x) P_\theta(dx) = g(\theta) \qquad (\forall \theta \in \Theta)$$

を満足するとき δ は $g(\theta)$ の**不偏推定量** (unbiased estimator) であるという．

今扱っている統計モデルでは，$\mathcal{P}$ としては $(\mathcal{X}, \mathcal{B}) = (\mathbf{R}^n, \mathbf{B}_n)$ 上の確率分布で，平均がある $\beta \in \mathbf{R}^p$ に対して $X\beta$ となり，分散共分散行列がある $\sigma^2 > 0$ に対して $\sigma^2 I_n$ となるもの全体を考えているのである．θ は，(β, σ^2) ではなく，$P \in \mathcal{P}$ そのものと考えられる．

ある確率空間に構成した確率変数 Y と $e = Y - X\beta$ は分布族の表現の便宜のためであって，すべての計算を $(\mathbf{R}^n, \mathbf{B}_n)$ 上で行うことは可能である．

「真のパラメータが (β, σ^2) のとき」というのは,「平均 $X\beta$，分散共分散行列 $\sigma^2 I_n$ をもつ確率測度 P_θ のもとで」という意味である．同様に「真のパラメータが β のとき」というのは,「平均が $X\beta$ で，ある $\sigma^2 > 0$ に対して分散共分散行列が $\sigma^2 I_n$ となる確率測度 P_θ のもとで」という意味である．β, σ^2 が真のパラメータであるときの期待値作用素を E_{β,σ^2} と書いた．これは Y の可測関数 $f(Y)$ に対して，

$$E_{\beta,\sigma^2}[f(Y)] = \int_{\mathbf{R}^n} f(y) P_\theta(dy) \tag{2.12}$$

と表されていることを意味する．ただし，式 (2.12) は，平均分散構造が $X\beta, \sigma^2 I_n$ となる P_θ に依存している．したがって，関数 f のクラスを適当に制限した場合に限り，左辺が意味をもつ．

定義 6. $c \in \mathbf{R}^p$ とする．$c \in L[X']$ のとき β の線形関数 $\mathbf{R}^p \ni \beta \mapsto c'\beta \in \mathbf{R}$ は**推定可能** (estimable) であるという．

定理 11. (a) $a \in \mathbf{R}^n$, $c \in \mathbf{R}^p$ とする．確率変数 Y の線形関数 $a'Y$ が β の線形関数 $c'\beta$ $(\beta \in \mathbf{R}^p)$ の不偏推定量であるための必要十分条件は $X'a = c$ となることである．

(b) 線形関数 $c'\beta$ ($\beta \in \mathbf{R}^p$) の線形不偏推定量が存在するための必要十分条件は $c'\beta$ が推定可能になることである．

証明．(a): $a'Y$ が不偏推定量であれば，定義より，任意の $\beta \in \mathbf{R}^p$ に対して，$a'X\beta = E[a'Y] = c'\beta$，したがって，$X'a = c$ である．逆も明らかである．
(b): $c'\beta$ の線形不偏推定量 $a'Y$ が存在すれば (a) より $c \in L[X']$ である．逆にこのとき，ある $a \in \mathbf{R}^n$ に対して，$c = X'a$ となり，$a'Y$ が $c'\beta$ の線形不偏推定量になっている．□

定理 12. 線形関数 $c'\beta$ が推定可能とする．このとき，c と $c'\hat{\beta}$ は，ある $b \in \mathbf{R}^p$ に対して $c = X'Xb$, $c'\hat{\beta} = b'X'Y$ と表される．$c'\hat{\beta}$ は正規方程式 (2.10) の解 $\hat{\beta}$ の取り方によらず，$c'\beta$ の線形不偏推定量である．

証明．$c \in L[X'] = L[X'X]$ だからあるベクトル $b \in \mathbf{R}^p$ に対して $c = X'Xb$. このとき，$c'\hat{\beta} = b'X'X\hat{\beta} = b'X'Y$ だから，これは $\hat{\beta}$ の取り方に依らない Y の線形関数である．さらに，$X'(Xb) = c$ だから定理 11 によって $b'X'Y$ は $c'\beta$ の不偏推定量である．□

定理 13. β の線形関数 $c'\beta$ が推定可能とする．このとき，正規方程式の任意の解 $\hat{\beta}$ に対して，$c'\hat{\beta}$ は $c'\beta$ の**最良線形不偏推定量** (best linear unbiased estimator，BLUE) である．すなわち，任意の (β, σ^2) に対して，

$$\mathrm{Var}[c'\hat{\beta}] = \min\{\mathrm{Var}[a'Y]; a'Y \text{ は } c'\beta\text{の線形不偏推定量 }\}$$

となる．ここで，Var は真値が (β, σ^2) のときの分散を表す．

証明．$c'\beta$ は推定可能であるから，$c = X'Xb$, $c'\hat{\beta} = b'X'Y$ と表される (定理 12)．$a'Y$ を $c'\beta$ の線形不偏推定量とすると定理 11 より $X'a = c$ となり，したがって $a'Xb = c'b = b'X'Xb$. このとき，$\mathrm{Cov}[a'Y - c'\hat{\beta}, c'\hat{\beta}] = \mathrm{Cov}[a'Y - b'X'Y, b'X'Y] = \sigma^2(a'Xb - b'X'Xb) = 0$ であるから，

$$\mathrm{Var}\left[a'Y\right] = \mathrm{Var}\left[a'Y - c'\hat{\beta}\right] + \mathrm{Var}\left[c'\hat{\beta}\right] \tag{2.13}$$

となり証明が終わる．□

次の定理は $c'\beta$ の最良線形不偏推定量を特徴づける．

定理 14. 次の (a),(b),(c) は同値である:

(a) 線形統計量 $a'Y$ は線形関数 $c'\beta$ の最良線形不偏推定量である．
(b) $c'\beta$ は推定可能で，任意の真値 (β, σ^2) に対して

$$c'\hat{\beta} = a'Y \quad \text{a.s.} \tag{2.14}$$

(c) $X'a = c$ かつ $a \in L[X]$.

証明．(a)⇒(b): 定理 11 より $c'\beta$ は推定可能である．また，(2.13) より $\mathrm{Var}[a'Y - c'\hat{\beta}] = 0$ でなければならないから，不偏性によって $a'Y = c'\hat{\beta}$ a.s. である．
(b)⇒(a): 定理 13 より明らか．
(b)⇒(c): $c'\beta$ は推定可能だから定理 12 より $c'\hat{\beta}$，したがって，$a'Y$ は $c'\beta$ の不偏推定量であり，定理 11 によって $X'a = c$ である．また，定理 12 の $b \in \mathbf{R}^p$ をとると，$c'\hat{\beta} = b'X'Y$ となるから，$0 = c'\hat{\beta} - a'Y = b'X'Y - a'Y = (b'X' - a')Y$．この分散を考えて $a = Xb \in L[X]$ を得る．
(c)⇒(b): ある $b \in \mathbf{R}^p$ に対して $a = Xb$ と書けるので $c = X'a = X'Xb$. このとき，$c'\hat{\beta} = b'X'X\hat{\beta} = b'X'Y = a'Y$．□

$X'X$ が非退化のときは $\hat{\beta}$ が一意に決まりその分散共分散行列は $\sigma^2(X'X)^{-1}$ となる．つぎの命題は $X'X$ が退化するときも，あたかも $\sigma^2(X'X)^-$ が $\hat{\beta}$ の分散共分散行列であるかのような計算が結果として成り立つことを示している．もっとも，退化する場合は $\hat{\beta}$ は意味がないので，これはあくまで便法である．

命題 5. (a) $c_i \in L[X']$ $(i = 1, 2)$ ならば $c_1'(X'X)^-c_2$ は $(X'X)^-$ の取り方に依らない．

(b) $c_i'\beta$ $(i=1,2)$ が推定可能線形関数とする．このとき，それぞれの最良不偏推定量 $c_i'\hat{\beta}$ の共分散は，(β,σ^2) が真値のとき，

$$\mathrm{Cov}\left[c_1'\hat{\beta}, c_2'\hat{\beta}\right] = \sigma^2 c_1'(X'X)^- c_2$$

と表される．

証明．$c_i'\beta$ が推定可能ならば，定理 12 より p 次元列ベクトル b_i が存在して，$c_i = X'Xb_i$, $c_i'\hat{\beta} = b_i'X'Y$ $(i=1,2)$ と表される．このとき，$\mathrm{Cov}\left[c_1'\hat{\beta}, c_2'\hat{\beta}\right] = \mathrm{Cov}\left[b_1'X'Y, b_2'X'Y\right] = \sigma^2 b_1'X'Xb_2 = \sigma^2 b_1'X'X(X'X)^-X'Xb_2 = \sigma^2 c_1'(X'X)^- c_2$ となり (b) を得る．左辺は右辺の $(X'X)^-$ の選び方に無関係だから右辺もそうであり，(a) も示された．□

問 5. β の任意の線形関数 $c'\beta$ が推定可能であるための必要十分条件は $\mathrm{rank}\ X = p$ となることである．証明せよ．

いままでは β の推定を考えてきたが，つぎに σ^2 の推定を考えよう．$\hat{\beta}$ を正規方程式の任意の解とする．定理 10 より，$\hat{Y} := X\hat{\beta} \equiv P_X Y$ は $\hat{\beta}$ の取り方に依らず意味がある．$e = Y - \hat{Y}$ は**残差ベクトル**と呼ばれる．

定理 15. (a) $e = Y - \hat{Y} = (I_n - P_X)Y = (I_n - P_X)\epsilon$. とくに，$E[e]=0$, $\mathrm{Var}[e] = \sigma^2(I_n - P_X)$.

(b) $c'\beta$ が推定可能ならば，$\mathrm{Cov}[c'\hat{\beta}, e] = 0$.

(c) $R_0^2 = \min_{\beta\in\mathbf{R}^p} |Y - X\beta|^2$ は次のように表現される：

$$\begin{aligned}
R_0^2 &= |Y - X\hat{\beta}|^2 \\
&= Y'Y - Y'X\hat{\beta} = Y'Y - \hat{\beta}'X'X\hat{\beta} \\
&= Y'Y - Y'X(X'X)^-X'Y = Y'(I_n - P_X)Y \\
&= \epsilon'(I_n - P_X)\epsilon = e'e.
\end{aligned}$$

証明．(a): 定義から

$$\begin{aligned} e &\equiv Y - \hat{Y} \equiv Y - X\hat{\beta} \\ &= Y - P_X Y = (I_n - P_X)Y \quad (\text{定理 10 から}) \\ &= (I_n - P_X)(X\beta + \epsilon) \quad (\beta\text{が真値のとき}) \\ &= (I_n - P_X)\epsilon. \end{aligned}$$

(b): $c'\beta$ が推定可能だから，ある p 次元ベクトル b に対して，$c'\hat{\beta} = b'X'Y$ と書ける（定理 12）．したがって，

$$\begin{aligned} \mathrm{Cov}\left[c'\hat{\beta}, e\right] &= \mathrm{Cov}\left[b'X'Y, (I_n - P_X)Y\right] \\ &= \sigma^2 b'X'(I_n - P_X) \\ &= 0. \end{aligned}$$

(c): 容易にわかるように：

$$\begin{aligned} R_0^2 &= |Y - X\hat{\beta}|^2 \quad (\text{定理 10}) \\ &= Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= Y'Y - \hat{\beta}'X'X\hat{\beta} = Y'Y - Y'X\hat{\beta} \quad (\text{正規方程式から } X'Y = X'X\hat{\beta}) \\ &= Y'Y - Y'P_X Y \quad (\text{定理 10 から } X\hat{\beta} = P_X Y) \\ &= Y'(I_n - P_X)Y = Y'(I_n - X(X'X)^{-}X')Y \\ &= \epsilon'(I_n - P_X)Y = \epsilon'(I_n - P_X)\epsilon \quad ((\text{a}) \text{ から}) \\ &= \epsilon'(I_n - P_X)(I_n - P_X)\epsilon \\ &= e'e \quad ((\text{a}) \text{ から}). \quad \square \end{aligned}$$

定理 16. 任意の $\beta \in \mathbf{R}^p$, $\sigma^2 > 0$ に対して

$$E_{\beta,\sigma^2}\left[R_0^2\right] = \sigma^2(n - r)$$

となる．ここで $r = \mathrm{rank}\ X$．とくに $n > r$ のとき，

$$\hat{\sigma}^2 := \frac{R_0^2}{n - r}$$

は σ^2 の不偏推定量である：

$$E_{\beta,\sigma^2}\left[\hat{\sigma}^2\right] = \sigma^2 \qquad (\forall\beta \in \mathbf{R}^p, \sigma^2 > 0).$$

証明．E_{β,σ^2} を E と表す．定理 15 より $R_0^2 = \epsilon'(I_n - P_X)\epsilon$ だから，

$$\begin{aligned}
E\left[R_0^2\right] &= E\left[\epsilon'(I_n - P_X)\epsilon\right] \\
&= \mathrm{tr}\left\{(I_n - P_X)\mathrm{Var}[\epsilon]\right\} + E[\epsilon]'(I_n - P_X)E[\epsilon] \qquad (\text{命題}\mathbf{??}) \\
&= \sigma^2 \mathrm{tr}(I_n - P_X) \\
&= \sigma^2(n - r). \qquad \square
\end{aligned}$$

2.6 仮説検定

この節ではパラメータ β が $\Theta = \mathbf{R}^p$ のある部分集合 Θ_0 に属するかそうでないかを観測に基づいて判断する問題を考える．これは後に体系的に議論する検定論の基本的な例であり，応用上も重要である．

我々は誤差項の分布に関して正規性を仮定する．すなわち，ガウス・マルコフモデル (2.8) において，さらに，

$$\epsilon \sim N_n(0, \sigma^2 I_n)$$

を仮定する．[*1]以前のように $R_0^2 = \min_{\beta\in\mathbf{R}^p} |Y - X\beta|^2$ とする．さらに $r := \mathrm{rank}\ X < n$ と仮定する．

定理 17. 任意の真値 (β, σ^2) のもとで，R_0^2/σ^2 はカイ 2 乗分布 $\chi^2(n-r)$ に従う．

証明．仮定から $\epsilon/\sigma \sim N_n(0, I_n)$ であり，定理 15 から $R_0^2/\sigma^2 = (\epsilon/\sigma)'(I_n - P_X)(\epsilon/\sigma)$，さらに $\mathrm{rank}\ (I_n - P_X) = n - r$ であるから，定理 7 によって，

[*1] この仮定は人工的に思われるかもしれない．そのような読者は，誤差分布の楕円型分布への拡張やロバスト推測論によって，より自然なモデルに導かれるであろう．

$R_0^2/\sigma^2 \sim \chi^2(n-r)$ である．□

つぎにガウス・マルコフモデルの期待値ベクトル $X\beta$ に線形制約がある場合を考えよう．$q < p$ とし，$p \times q$ 行列 H が $L[H] \subset L[X']$，$\mathrm{rank}\, H = q$ を満足すると仮定する．d をある q 次元ベクトルとし，$\Theta_0 = \{\beta \in \mathbf{R}^p; H'\beta = d\}$ とする．H のランクの仮定から $\Theta_0 \neq \emptyset$ である．

$$R_1^2 = \min_{\beta \in \Theta_0} |Y - X\beta|^2$$

とおく．次の結果はパラメータの検定において基本的である．

定理 18. H, d が上の仮定を満たすとする．このとき，

(a) 任意の真値 β, σ^2 に対して，R_0^2 と $R_1^2 - R_0^2$ は独立であり，$R_0^2/\sigma^2 \sim \chi^2(n-r)$. さらに，ある $\delta \in \mathbf{R}_+$ に対して，$(R_1^2 - R_0^2)/\sigma^2 \sim \chi^2(q, \delta)$. とくに，

$$\frac{(R_1^2 - R_0^2)/q}{R_0^2/(n-r)} \sim F(q, n-r, \delta).$$

(b) 真値 β が $\beta \in \Theta_0$ を満足するとき，$(R_1^2 - R_0^2)/\sigma^2 \sim \chi^2(q)$ であり，とくに，

$$\frac{(R_1^2 - R_0^2)/q}{R_0^2/(n-r)} \sim F(q, n-r).$$

証明．パラメータの真値を β_0, σ_0^2 で表す．$H'\beta_1 = d$ なる $\beta_1 \in \mathbf{R}^p$ が存在する．$\gamma = \beta - \beta_1 \in \mathbf{R}^p$ によってパラメータ β をパラメータ γ に変換する．定理 17 より，任意の真値 β_0, σ_0^2 に対して，$R_0^2/\sigma_0^2 \sim \chi^2(n-r)$ である．また，定理 15 から，$R_0^2 = Y'(I_n - P_X)Y = (Y - X\beta_1)'(I_n - P_X)(Y - X\beta_1)$ となる．いっぽう，

$$\begin{aligned} R_1^2 &= \min_{\beta \in \Theta_0} |Y - X\beta|^2 = \min_{\gamma : H'\gamma = 0} |(Y - X\beta_1) - X\gamma|^2 \\ &= (Y - X\beta_1)'(I_n - P_{X|H})(Y - X\beta_1) \end{aligned}$$

であるが，$I_n - P_{X|H}$ がべき等，対称であるから，定理 7 によって，$R_1^2/\sigma_0^2 \sim \chi^2(m,\delta)$ であることがわかる．ここで，

$$\delta = \sigma_0^{-2}(X\beta_0 - X\beta_1)'(I_n - P_{X|H})(X\beta_0 - X\beta_1),$$

m は自由度であるが，定理 2 から $m = n - \dim\,(L[X|H]) = n - r + q$ である．

定義から $R_1^2 - R_0^2 \geq 0$ は自明であり，2 次形式の分解 $R_1^2 = R_0^2 + (R_1^2 - R_0^2)$ に対して定理 9 を適用すると，$(R_1^2 - R_0^2)/\sigma_0^2$ は R_0^2 と独立であり，$\chi^2(q,\delta)$ に従うことがわかる．

とくに真値 $\beta_0 \in \Theta_0$ のとき，$H'(\beta_0 - \beta_1) = 0$ だから $X(\beta_0 - \beta_1) \in L[X|H]$，したがって，$(I_n - P_{X|H})(X\beta_0 - X\beta_1) = 0$ となり，$\delta = 0$ である．□

命題 6. $\hat{\beta}$ を正規方程式の任意の解とするとき，

$$R_1^2 - R_0^2 = (H'\hat{\beta} - d)'\left\{H'(X'X)^- H\right\}^{-1}(H'\hat{\beta} - d).$$

証明. 定理 18 の証明の β_1 と定理 2 の G をとる．$H'\hat{\beta} = G'X'X(X'X)^- XY = G'X'Y$, $d = H'\beta_1 = G'X'X\beta_1$ だから，$G'X'(Y - X\beta_1) = H'\hat{\beta} - d$. また，$H'(X'X)^- H = G'X'X(X'X)^- X'XG = G'X'XG$ だから，定理 18 の証明の R_1^2 の表現と定理 2 から，

$$\begin{aligned} R_1^2 - R_0^2 &= (Y - X\beta_1)'XG(G'X'XG)^{-1}G'X'(Y - X\beta_1) \\ &= (H'\hat{\beta} - d)'\left\{H'(X'X)^- H\right\}^{-1}(H'\hat{\beta} - d) \end{aligned}$$

となる．□

この命題によると $R_1^2 - R_0^2$ は $X, H, d, \hat{\beta}$ で計算できる．また，推定可能関数のベクトル $H'\beta$ の推定量 $H'\hat{\beta}$ の分散は $\mathrm{Var}[H'\hat{\beta}] = \sigma^2 H'(X'X)^- H$ であった (命題 5)．$(R_1^2 - R_0^2)/\sigma^2$ の表現は $H'\hat{\beta} - d$ の 2 乗がこの分散で規格化されている．さらに，真値 $\beta_0 \in \Theta_0$ のとき，$H'\hat{\beta} - d$ の平均は零になるので，$(R_1^2 - R_0^2)/\sigma^2$ が（中心）カイ 2 乗分布になることがこの表現からもわかる．

さて，上述の結果を検定問題

$$H_0:\ H'\beta = d \quad \text{vs.} \quad H_1:\ H'\beta \neq d \tag{2.15}$$

に適用しよう．仮説検定は第３章で詳しく議論されるので，ここでは直観的な説明をするにとどめる．まず，仮説 H_0, H_1 はそれぞれ $\beta \in \Theta_0$, $\beta \in \mathbf{R}^p - \Theta_0$ の意味である．仮説 H_0 は**帰無仮説**，仮説 H_1 は**対立仮説**と呼ばれ，我々はデータをもとに帰無仮説が対立仮説と比して統計的に妥当なものか判断する．検定問題 (2.15) に対しては，データから計算した

$$F := \frac{(R_1^2 - R_0^2)/q}{R_0^2/(n-r)}$$

の値が大きいとき，H_0 を棄却する．R_0^2 の分布は H_0 の下でも，H_1 の下でも同じであるが，$(R_1^2 - R_0^2)/\sigma^2$ は非心 χ^2 分布に従うので，H_1 に対しては H_0 におけるよりも大きい値が出やすい傾向がある．このことは，たとえば $\chi^2(q,\delta)$ の平均値が $q+\delta$ であることからも理解できる．

データの採取以前にある確率の値 $\alpha \in (0,1)$ を設定し，f_α を F 分布 $F(q, n-r)$ の上側 $100\alpha\%$ 点とし，$F \geq f_\alpha$ であるとき F が「大きい」と判断される．α を**有意水準**と呼び，通常 $\alpha = 0.05$ や $\alpha = 0.01$ などが用いられる．帰無仮説 H_0 が正しいという仮定の下で $F \geq f_\alpha$ となる確率は α しかないにもかかわらず，実際にそれが起きたとすると，仮説 H_0 は疑わしく，これを棄却するのである．F 分布に基づく検定は F **検定** (F-test) と呼ばれる．

このように，データのなすベクトルの２次形式による，データ変動の分解をもとに検定を構成する方法を**分散分析** (analysis of variance) と呼ぶ．分散分析を行うには**分散分析表** (analysis-of-variance table, ANOVA table) を作ると便利である．

	平方和	自由度	平均平方和	F
仮説による変動	$R_1^2 - R_0^2$	q	$(R_1^2 - R_0^2)/q$	F
残差による変動	R_0^2	$n-r$	$R_0^2/(n-r)$	
全変動	R_1^2	$n-(r-q)$		

分散分析表にはデータから計算した数値を入れる．また，１つの平方和は他

の 2 つから求めてよい．

2.7 平均の検定

平均 μ および分散 σ^2 が未知の正規分布 $N(\mu, \sigma^2)$ からの無作為標本 Y_j $(j = 1, ..., n)$ を観測するとし，パラメータ μ に関する仮説 H_0: $\mu = \mu_0$ を対立仮説 H_1: $\mu \neq \mu_0$ に対して検定する問題を考える．ガウス・マルコフモデルを $X = [1, ..., 1]'$, $\beta = \mu$, $H = 1$, $d = \mu_0$ に対して適用すると，検定統計量 $F = n(\bar{Y} - \mu_0)^2/(\sum_{j=1}^n (Y_j - \bar{Y})^2/(n-1))$ が大きいとき H_0 を棄却する検定となる．ここで $\bar{Y} = n^{-1}\sum_{j=1}^n Y_j$．あるいは同じことだが，$T = \sqrt{n}(\bar{Y} - \mu_0)/\sqrt{\sum_{j=1}^n (Y_j - \bar{Y})^2/(n-1)}$ の絶対値が大きいとき H_0 を棄却するといってもよく，これは**両側 t 検定**になる．すなわち T は H_0 のもとで自由度 $n-1$ の t 分布に従い，これを基に構成された棄却域 $|t| \geq t_\alpha(n-1)$ は実軸上左右両側になる．ここで $t_\alpha(\nu)$ は t 分布 $t(\nu)$ の上側 $100\alpha/2\%$ 点である．

つぎに 2 標本の場合を考える．$N(\mu_1, \sigma^2)$ からの無作為標本 Y_{1j} $(j = 1, ..., n_1)$ と，それと独立な $N(\mu_2, \sigma^2)$ からの無作為標本 Y_{2l} $(l = 1, ..., n_2)$ を観測するものとする．μ_1, μ_2, σ^2 はすべて未知とし，仮説 H_0: $\mu_1 = \mu_2$ を対立仮説 H_1: $\mu_1 \neq \mu_2$ に対して検定する．ガウス・マルコフモデルを適用することによって F 検定統計量が得られる．同値な t 検定は統計量

$$T = \frac{(\bar{Y}_1 - \bar{Y}_2)/\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}{\sqrt{\left[\sum_{j=1}^{n_1}(Y_{1j} - \bar{Y}_1)^2 + \sum_{l=1}^{n_2}(Y_{2l} - \bar{Y}_2)^2\right]/(n_1 + n_2 - 2)}}$$

で与えられる．ここで，$\bar{Y}_1 = n_1^{-1}\sum_{j=1}^{n_1} Y_{1j}$, $\bar{Y}_2 = n_2^{-1}\sum_{l=1}^{n_2} Y_{2l}$ である．T は帰無仮説のもとで，自由度 $n_1 + n_2 - 2$ の t 分布に従い，$|T|$ が大きいとき H_0 を棄却する．

2.8　重回帰分析

重回帰分析においては切片項を含むガウス・マルコフモデルが用いられる．すなわち，変数 y を変数 $x_1, ..., x_p$ と定数項で説明するために線形モデル

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \varepsilon$$

を考える（**重回帰モデル**）．ここで，右辺の β_0 が定数項を表している．n 個の観測がある場合，モデルは

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i \qquad (i = 1, 2, ..., n)$$

と表される．i が異なる個体の番号を表すとき，誤差項の独立性が成り立つであろう．このようにして，我々のモデルはガウス・マルコフモデル

$$Y = X\beta + \epsilon$$

で表現される．ここで，

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \qquad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \qquad \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix},$$

$$X = [\mathbf{1}|X_1] \quad (n \times (p+1)), \qquad \mathbf{1} = [1, ..., 1]' \quad (n \times 1), \qquad X_1 = [x_{ia}] \quad (n \times p),$$

$$E[\epsilon] = 0, \qquad \mathrm{Var}[\epsilon] = \sigma^2 I_n$$

である．

$(p+1)$ 次正方行列 $X'X$ は

$$X'X = \begin{bmatrix} n & \mathbf{1}'X_1 \\ X_1'\mathbf{1} & X_1'X_1 \end{bmatrix}$$

となり，$\tilde{\beta} = [\beta_1, ..., \beta_p]'$ と書くとき，正規方程式 $X'X\beta = X'Y$ は

$$n\beta_0 + \mathbf{1}'X_1\tilde{\beta} = \mathbf{1}'Y \tag{2.16}$$

$$X_1'\mathbf{1}\beta_0 + X_1'X_1\tilde{\beta} = X_1'Y \tag{2.17}$$

となる．

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \qquad \bar{x}_a = \frac{1}{n}\sum_{i=1}^{n} x_{ia}$$

とおくとき，(2.16) から

$$\beta_0 = \bar{Y} - \frac{1}{n}\mathbf{1}'X_1\tilde{\beta} = \bar{Y} - \sum_{a=1}^{p} \bar{x}_a\beta_a \tag{2.18}$$

となる．また，(2.17) から

$$X_1'\left(\mathbf{1}\beta_0 + X_1\tilde{\beta} - \bar{Y}\mathbf{1}\right) = X_1'\left(Y - \bar{Y}\mathbf{1}\right) \tag{2.19}$$

だが,

$$\begin{aligned}\mathbf{1}\beta_0 + X_1\tilde{\beta} - \bar{Y}\mathbf{1} &= X_1\tilde{\beta} - \frac{1}{n}\mathbf{1}\mathbf{1}'X_1\tilde{\beta} \qquad ((2.18)\text{ より})\\ &= \tilde{X}\tilde{\beta}.\end{aligned} \tag{2.20}$$

ここで,

$$\tilde{X} = X_1 - \frac{1}{n}\mathbf{1}\mathbf{1}'X_1 = [x_{ia} - \bar{x}_a].$$

容易にわかるように,

$$\tilde{X}'\tilde{X} = X_1'\tilde{X}. \tag{2.21}$$

さらに，$\mathbf{1}'(Y - \bar{Y}\mathbf{1}) = 0$ から,

$$X_1'(Y - \bar{Y}\mathbf{1}) = (X_1 - \frac{1}{n}\mathbf{1}\mathbf{1}'X_1)'(Y - \bar{Y}\mathbf{1}) = \tilde{X}'(Y - \bar{Y}\mathbf{1}). \tag{2.22}$$

(2.19)-(2.22) から

$$\tilde{X}'\tilde{X}\tilde{\beta} = \tilde{X}'(Y - \bar{Y}\mathbf{1}) \tag{2.23}$$

となる．係数 s_{ab}, s_{ya} を

$$s_{ab} = \frac{1}{n}\sum_{i=1}^{n}(x_{ia} - \bar{x}_a)(x_{ib} - \bar{x}_b) \qquad (a, b = 1, ..., p)$$

$$s_{ya} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(x_{ia} - \bar{x}_a) \qquad (a = 1, ..., p)$$

とおくと，結局，(2.18) と (2.23) から正規方程式は，

$$\begin{cases} \sum_{b=1}^{p} s_{ab}\beta_b = s_{ya} \qquad (a = 1, ..., p) \\ \\ \beta_0 = \bar{Y} - \sum_{a=1}^{p} \bar{x}_a\beta_a \end{cases} \tag{2.24}$$

となる．行列 $s = [s_{ab}]_{a,b=1}^{p}$ が正則であると仮定し，逆行列を $s^{-1} = [s^{ab}]$ で表すと，(2.24) の解 $\hat{\beta}$ は

$$\begin{cases} \hat{\beta}_a = \sum_{b=1}^{p} s^{ab}s_{yb} \qquad (a = 1, ..., p) \\ \\ \hat{\beta}_0 = \bar{Y} - \sum_{a=1}^{p} \bar{x}_a\hat{\beta}_a \end{cases} \tag{2.25}$$

となる．

σ^2 の推定量 $\hat{\sigma}^2$ は $\hat{\sigma}^2 = R_0^2/(n - p - 1)$ である．

以後誤差分布は $\epsilon \sim N_n(0, \sigma^2 I_n)$ であると仮定し，幾つかの仮説検定問題を考察する．命題 6 を適用するためには $(X'X)^{-1}$ の表現が必要となる．命題 3(d) と

$$X_1'X_1 - \frac{1}{n}X_1'\mathbf{1}\mathbf{1}'X_1 = \tilde{X}'\tilde{X} = ns$$

を用いて，

$$(X'X)^{-1} = \frac{1}{n}\begin{bmatrix} 1 + \frac{1}{n}\mathbf{1}'X_1 s^{-1}\frac{1}{n}X_1'\mathbf{1} & -\frac{1}{n}\mathbf{1}'X_1 s^{-1} \\ -s^{-1}\frac{1}{n}X_1'\mathbf{1} & s^{-1} \end{bmatrix}$$

を得る．

$p=1$ の場合，重回帰モデルは**単回帰モデル**

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \qquad (i=1,...,n)$$

となり，$\hat{\beta}_1 = s_{yx}/s_{xx}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ となる．ここで，$\bar{x} = n^{-1}\sum_{i=1}^n x_i$, $\bar{y} = n^{-1}\sum_{i=1}^n y_i$, $s_{xx} = n^{-1}\sum_{i=1}^n (x_i-\bar{x})^2$, $s_{yx} = n^{-1}\sum_{i=1}^n (y_i-\bar{y})(x_i-\bar{x})$ である．

a. $\beta_a = 0$ の検定

$a \in \{1,...,p\}$ に対して，検定問題

$$H_0 : \beta_a = 0 \quad \text{vs.} \quad H_1 : \beta_a \neq 0$$

を考えよう．これは変数 x_a が y の説明に必要かどうかの検定である．$(p+1)$ 次元ベクトル

$$H = [0 \cdots 0\ 1\ 0 \cdots 0]' \qquad (a \text{ に対応する成分のみ } 1)$$

と $d=0$ に対して命題 6 を適用すると，

$$R_1^2 - R_0^2 = \frac{\hat{\beta}_a^2}{s^{aa}/n}$$

だから，定理 18 によって，

$$\frac{\hat{\beta}_a^2}{\hat{\sigma}^2 s^{aa}/n} \sim F(1, n-p-1) \qquad (H_0\text{のもとで})$$

となる．対立仮説 H_1 のもとではこの統計量は非心 F 分布に従い，値は大きくでる傾向がある．そこで，我々は

$$\frac{\hat{\beta}_a^2}{\hat{\sigma}^2 s^{aa}/n} \geq F_\alpha(1, n-p-1)$$

ならば H_0 を棄却し，そうでなければ H_0 を採択する．ここで，$F_\alpha(1,n-p-1)$ は F 分布の上側 $100\alpha\%$ 点で，$\alpha \in (0,1)$ は有意水準である．

また，同じことであるが，t 検定

$$\frac{|\hat{\beta}_a|}{\sqrt{\hat{\sigma}^2 s^{aa}/n}} \geq t_\alpha(n-p-1) \Rightarrow H_0\text{を棄却}$$

を用いてもよい．ここで，$t_\alpha(\nu)$ は自由度 ν の t 分布の上側 $100\alpha/2\%$ 点である．

b. $\beta_0 = 0$ **の検定**

切片項が y の説明に必要か否かを検定する．検定問題

$$H_0 : \beta_0 = 0 \quad \text{vs.} \quad H_1 : \beta_0 \neq 0$$

を考える．今度は $(p+1)$ 次元ベクトル $H = [1\ 0 \cdots 0]'$ と $d = 0$ に対して，命題 6 と定理 18 と適用する．

$$\begin{aligned} H'(X'X)^{-1}H &= \frac{1}{n}\left(1 + \frac{1}{n}\mathbf{1}'X_1 s^{-1}\frac{1}{n}X_1'\mathbf{1}\right) \\ &= \frac{1}{n}\left(1 + \sum_{a,b=1}^{p} \bar{x}_a \bar{x}_b s^{ab}\right) \end{aligned}$$

に注意すると，

$$\frac{\hat{\beta}_0^2}{\hat{\sigma}^2\left(1 + \sum_{a,b=1}^{p} \bar{x}_a \bar{x}_b s^{ab}\right)/n} \sim F(1, n-p-1) \qquad (H_0\text{のもとで})$$

となるから，

$$\frac{|\hat{\beta}_0|}{\sqrt{\hat{\sigma}^2\left(1 + \sum_{a,b=1}^{p} \bar{x}_a \bar{x}_b s^{ab}\right)/n}} \geq t_\alpha(n-p-1) \Rightarrow H_0\text{を棄却}$$

とすればよい．

c. 回帰の有意性

説明変数 $x_1, ..., x_p$ が y の説明に全体として，役立つか否か検定する．検定問題

$$H_0 : \beta_1 = \cdots = \beta_p = 0 \quad \text{vs.} \quad H_1 : \text{そうでない}$$

を考えよう．

今の場合制約条件を表す行列 H は $H' = [0 \mid I_p]$ $(p \times (p+1))$，$d = 0$ となる．$\mathbf{1} \in L[X]$ から容易にわかる分解

$$|Y - \bar{Y}\mathbf{1}|^2 = |Y - X\hat{\beta}|^2 + |X\hat{\beta} - \bar{Y}\mathbf{1}|^2$$

によって，$R_1^2 - R_0^2 = |X\hat{\beta} - \bar{Y}\mathbf{1}|^2$ となる．検定統計量は

$$F = \frac{|X\hat{\beta} - \bar{Y}\mathbf{1}|^2/p}{\hat{\sigma}^2},$$

帰無分布は $F(p, n-p-1)$ である．F の値が大きいとき H_0 を棄却し，回帰が有意であると判断される．

$R^2 = |X\hat{\beta} - \bar{Y}\mathbf{1}|^2/|Y - \bar{Y}\mathbf{1}|^2$ は**決定係数**と呼ばれ，データへの回帰モデルの当てはまりの度合いを表すと考えられる．

例 2. 3 種類の薬品 A_1, A_2, A_3 を調合して作った薬の効果を調べた．30 回実験を行い，各回の効果 Y と使用した薬品 A_1, A_2, A_3 の量 X_1, X_2, X_3 を表にまとめた．

Y	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y	X_1	X_2	X_3
66.3	2.62	7.13	9.43	72.8	2.98	8.66	7.77	75.8	2.95	8.21	8.72
74.6	3.22	7.89	8.52	64.2	2.79	8.31	5.54	61.3	3.17	6.47	7.21
59.6	2.37	6.38	8.17	59.5	3.12	6.51	7.13	61.6	2.54	7.27	7.02
64.8	3.32	7.68	6.60	70.4	3.14	7.36	8.46	57.5	2.34	5.68	6.55
61.9	3.36	7.27	6.23	62.0	2.75	6.65	8.42	62.8	2.72	7.46	7.44
68.9	3.05	7.85	8.36	70.7	3.48	7.18	7.37	70.9	3.73	7.98	7.43
73.6	3.17	9.07	7.17	55.1	2.79	5.39	7.06	66.0	2.89	7.38	7.25
71.3	2.82	7.86	7.45	67.2	2.31	7.82	8.61	70.3	3.62	7.57	7.82
66.1	3.04	6.78	6.78	70.4	3.04	8.81	8.20	66.4	3.38	8.07	7.06
56.7	2.76	6.28	7.58	60.7	2.61	7.99	5.44	76.8	2.71	8.75	6.92

データに基づき，パラメータの推定と分散分析を行い，次の表の結果を得た．p 値すなわち 33.09 より大きな F が出る確率は 0.0001 より小さく，回

帰は有意であり，決定係数は $R^2 = 0.792$ である．

β	推定値	t 値	p 値
β_0	3.41	0.49	0.630
β_1	3.53	2.35	0.027
β_2	4.75	7.86	0.000
β_3	2.27	3.94	0.001

	平方和	自由度	平均平方和	F
回帰による変動	799.5	3	266.5	33.09
残差による変動	209.4	26	8.054	
全変動	1008.9	29		

d.　予測

観測 $Y_1, ..., Y_n$ に基づいてモデルを当てはめ，説明変数 $x_1, ..., x_p$ のある値 $\xi_1, ..., \xi_p$ に対して，新たな観測 $Y_\xi = \beta_0 + \beta_1\xi_1 + \cdots + \beta_p\xi_p + \epsilon_*$ の予測をする．ここで，$\xi = [\xi_1, ..., \xi_p]'$，また，$\epsilon_*$ は $\epsilon_1, ..., \epsilon_n$ と同じ分布に従う独立なノイズである．Y_ξ の予測値は

$$\hat{Y}_\xi = \hat{\beta}_0 + \hat{\beta}_1\xi_1 + \cdots + \hat{\beta}_p\xi_p = c'_\xi\hat{\beta}$$

で与えられる．ここで $c_\xi = [1\ \xi']$ である．予測値 $\hat{Y}_\xi$ はデータ Y に依存する確率変動を含んだ量である．パラメータの真値 β に対して $\mu_\xi = c'_\xi\beta$ とするとき，$\hat{Y}_\xi \sim N(\mu_\xi, \sigma^2 c'_\xi(X'X)^{-1}c_\xi)$．また，定理 15(b) と正規性から η と $\hat{\sigma}^2 = e'e/(n-p-1)$ は独立になる．さらに，Y_ξ はこれらの変数と独立で，$Y_\xi \sim N(\mu_\xi, \sigma^2)$ である．

$$c'_\xi(X'X)^{-1}c_\xi = \frac{1}{n}\left\{1 + (\xi - \bar{x})'s^{-1}(\xi - \bar{x})\right\}, \qquad \bar{x} := [\bar{x}_a]$$

だから，

$$(Y_\xi - \hat{Y}_\xi)\Big/\sqrt{\left[1 + \frac{1}{n}\{1 + (\xi - \bar{x})'s^{-1}(\xi - \bar{x})\}\right]\hat{\sigma}^2} \sim t(n-p-1)$$

となる．したがって，$\alpha \in (0,1)$ に対して，t 分布 $t(n-p-1)$ の上側 $100\alpha/2\%$ 点を $t_\alpha(n-p-1)$ とするとき，Y_ξ が入るであろう信頼係数 $1-\alpha$ の**予測域**が区間

$$\Big[\hat{Y}_\xi - t_\alpha(n-p-1)\hat{\sigma}\sqrt{1+\frac{1}{n}\left\{1+(\xi-\bar{x})'s^{-1}(\xi-\bar{x})\right\}},$$
$$\hat{Y}_\xi + t_\alpha(n-p-1)\hat{\sigma}\sqrt{1+\frac{1}{n}\left\{1+(\xi-\bar{x})'s^{-1}(\xi-\bar{x})\right\}}\Big]$$

で与えられる．

e. μ_ξ の信頼区間

上と同じ記号を用いると，

$$(\hat{Y}_\xi - \mu_\xi)\Big/\sqrt{\frac{1}{n}\left\{1+(\xi-\bar{x})'s^{-1}(\xi-\bar{x})\right\}\hat{\sigma}^2} \sim t(n-p-1)$$

であるから，区間

$$\Big[\hat{Y}_\xi - t_\alpha(n-p-1)\hat{\sigma}\sqrt{\frac{1}{n}\left\{1+(\xi-\bar{x})'s^{-1}(\xi-\bar{x})\right\}},$$
$$\hat{Y}_\xi + t_\alpha(n-p-1)\hat{\sigma}\sqrt{\frac{1}{n}\left\{1+(\xi-\bar{x})'s^{-1}(\xi-\bar{x})\right\}}\Big]$$

がパラメータの真値 μ_ξ を覆う確率は $1-\alpha$ になる．この区間をパラメータ μ_ξ の，信頼係数 $1-\alpha$ の**信頼区間**と呼ぶ．

2.9 一元配置

因子 A に水準 $A_1, A_2, ..., A_p$ があり，水準 A_i での因子 A の効果を μ_i で表すとき，水準 A_i に対する j 番目の観測が

$$Y_{ij} = \quad \mu_i + \epsilon_{ij} \qquad (i=1,...,p;\ j=1,...,n_i) \tag{2.26}$$

という構造で規定されるとする．ここで，ϵ_{ij} $(i=1,...,p;\ j=1,...,n_i)$ は誤差を表す独立な確率変数で，以後 ϵ_{ij} は正規分布 $N(0,\sigma^2)$ に従うと仮定する．因子 A の水準 $A_1, A_2, ..., A_p$ における効果に差があるかどうか判断する

ために，検定問題

$$H_0:\ \mu_1 = \cdots = \mu_p \text{ vs. } H_1:\ \text{ある } i,j \text{ に対して} \mu_i \neq \mu_j \quad (2.27)$$

を考える．

$n = \sum_{i=1}^{p} n_i$ とし，n 次元確率ベクトル Y と ϵ を

$$Y = [Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, ..., Y_{p1}, ..., Y_{pn_p}]',$$
$$\epsilon = [\epsilon_{11}, ..., \epsilon_{1n_1}, \epsilon_{21}, ..., \epsilon_{2n_2}, ..., \epsilon_{p1}, ..., \epsilon_{pn_p}]'$$

とする．$\mathbf{1}_k = [1, ..., 1]' \in \mathbf{R}^k$ とし，$n \times p$ 行列 X と p 次元ベクトル β を

$$X = \begin{bmatrix} \mathbf{1}_{n_1} & & & O \\ & \mathbf{1}_{n_2} & & \\ & & \ddots & \\ O & & & \mathbf{1}_{n_p} \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

とすれば，(2.26) はガウス・マルコフモデル $Y = X\beta + \epsilon$ として表される．線形制約を表す $p \times (p-1)$ 行列 H を

$$H' = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & O & \\ & O & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix}$$

と取ると，(2.27) は検定問題 (2.15) と同じ形になる．今の場合，

$$|Y - X\beta|^2 = \sum_{ij} (Y_{ij} - \mu_i)^2$$

であるから，自由にパラメータを動かしたときの誤差 R_0^2 と制約条件のもとでの誤差 R_1^2 は

$$R_0^2 = \sum_{ij} (Y_{ij} - \bar{Y}_{i\cdot})^2, \quad R_1^2 = \sum_{ij} (Y_{ij} - \bar{Y}_{\cdot\cdot})^2$$

となる．ここで，

$$\bar{Y}_{i\cdot} = \sum_j Y_{ij}/n_i, \qquad \bar{Y}_{\cdot\cdot} = \sum_{ij} Y_{ij}/n$$

である．さらに，分解

$$\sum_{ij}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2 = \sum_i n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 + \sum_{ij}(Y_{ij} - \bar{Y}_{i\cdot})^2$$

より

$$R_1^2 - R_0^2 = \sum_i n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

であることがわかる．R_0^2 を級内変動，$R_1^2 - R_0^2$ を級間変動と呼ぶ．rank $X = p$, rank $H = p - 1$ だから，検定統計量は

$$F = \frac{\sum_i n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2/(p-1)}{\sum_{ij}(Y_{ij} - \bar{Y}_{i\cdot})^2/(n-p)}$$

である．正規性の仮定から，帰無仮説のもとで F は F 分布 $F(p-1, n-p)$ に従う．

一元配置のモデルは，因子 A の平均効果を μ で表し，平均 μ を基準とした A_i における相対的な効果を α_i で表すこともある．すなわち，

$$\mu_i = \mu + \alpha_i, \qquad \mu = \frac{1}{n}\sum_{i=1}^{p} n_i\mu_i$$

とするのである．このとき，

$$\sum_{i=1}^{p} n_i\alpha_i = 0$$

であるから，仮説 H_0 は

$$\alpha_1 = \cdots = \alpha_p = 0$$

と同じことである．

例 3. 4 種類の薬品 A_1, A_2, A_3, A_4 をラットに与え，その効果を調べたところ，表のような結果が得られた.

A_1	A_2	A_3	A_4
4.25	4.12	4.08	4.44
3.85	4.39	4.51	4.29
3.71	4.11	4.28	4.15
4.05	3.98	4.17	4.60
4.17	4.03	4.23	-
-	3.74	4.27	-

ラットの条件は均一であるとするとき，薬品 A_1, A_2, A_3, A_4 の効果の間に差があるかどうかを有意水準を $\alpha = 0.05$ で検定する．F 値は $F = 3.6095$ である．F 分布 $F(3, 17)$ の上側 5% 点は数表から $f_{0.05} = 3.20$ であるので，仮説 H_0 は棄却される．つまり薬品の間に差があると結論される．統計パッケージを用いると，$F > 3.6095$ となる確率 0.0350 が表示されるので $f_{0.05}$ の値を調べる必要はないし，今のデータならば数値をそのまま打ち込んでも手間はそれほどかからない．もし手計算で F を求めるのであれば，F は 1 次関数による変換で不変であるので，データから適当な共通の定数を引いたり，零でない定数を掛けたりして，計算しやすい数値に変換してから計算するとよい.

要因	平方和	自由度	平均平方和	F
級間	0.4088919	3	0.136297	3.6095
級内	0.6419367	17	0.037761	
全変動	1.0508286	20		

2.10 二元配置

２つの因子 A, B があり，A の水準 $A_1, ..., A_a$ と B の水準 $B_1, ..., B_b$ に対して観測値 Y_{ij} $(i = 1, ..., a;\ j = 1, ..., b)$ が得られているとする．観測される確率変数 Y_{ij} の構造が

$$Y_{ij} = \mu^* + \alpha_i^* + \beta_j^* + \epsilon_{ij} \tag{2.28}$$

であるとする．ここで，μ^* は水準によらない平均を表し，α_i^* は A の水準 A_i における効果を，β_j^* は B の水準 B_j における効果を表す．ϵ_{ij} は誤差項で独立に $N(0, \sigma^2)$ に従うと仮定する．(2.28) の $(\mu^* + \alpha_i^* + \beta_j^*)$ は $(E[Y_{ij}])$ の表現として冗長であるので，通常パラメータに制約条件

$$\sum_i \alpha_i^* = \sum_j \beta_j^* = 0$$

を置く．

二元配置のモデル (2.28) が表現している分布族はモデル

$$Y_{ij} = \alpha_i + \beta_j + \epsilon_{ij} \tag{2.29}$$

のものと同じである．実際，$\bar{\alpha}. = \sum_i \alpha_i / a$, $\bar{\beta}. = \sum_j \beta_j / b$ と定め，

$$\mu^* = \bar{\alpha}. + \bar{\beta}.,\ \alpha_i^* = \alpha_i - \bar{\alpha}.,\ \beta_j^* = \beta_j - \bar{\beta}.$$

とすれば，モデル (2.29) はモデル (2.28) に含まれていることがわかるし，逆に

$$\alpha_i = \alpha_i^*,\ \beta_j = \mu^* + \beta_j^*$$

とすれば，モデル (2.28) がモデル (2.29) に含まれていることがわかる．ただし，分布のパラメータ表示が一意でないので，パラメータ間に一対一対応はつかない．

$A_1, ..., A_a$ の効果に差があるかどうかの検定は

$$H_0' : \alpha_1 = \cdots = \alpha_a \ \text{ vs. } \ H_1' : \text{ある } i_1, i_2 \text{に対して} \alpha_{i_1} \neq \alpha_{i_2} \tag{2.30}$$

となる．ちなみに，H_0' は $E[Y_{ij}]$ が j にのみ依存するということで，モデル (2.28) の表現では，このことは仮説 $\alpha_1^* = \cdots = \alpha_a^* = 0$ と同値である．

さて，

$$Y = [Y_{11}, ..., Y_{1b}, Y_{21}, ..., Y_{2b}, ..., Y_{a1}, ..., Y_{ab}]',$$
$$\epsilon = [\epsilon_{11}, ..., \epsilon_{1b}, \epsilon_{21}, ..., \epsilon_{2b}, ..., \epsilon_{a1}, ..., \epsilon_{ab}]'$$

と表し，$ab \times (a+b)$ 行列 X と $a+b$ 次元パラメータ β を

$$X = \begin{bmatrix} \mathbf{1}_b & & & O & | & I_b \\ & \mathbf{1}_b & & & | & I_b \\ & & \ddots & & | & \vdots \\ & O & & \mathbf{1}_b & | & I_b \end{bmatrix}, \quad \beta = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_a \\ \beta_1 \\ \vdots \\ \beta_b \end{bmatrix}$$

とすれば，(2.29) はガウス・マルコフモデル $Y = X\beta + \epsilon$ になる．

線形制約を表す $(a+b) \times (a-1)$ 行列 H を

$$H' = \begin{bmatrix} 1 & -1 & & & & | & 0 & \cdots & 0 \\ & 1 & -1 & O & & | & \vdots & \cdots & \vdots \\ & O & \ddots & \ddots & & | & \vdots & \cdots & \vdots \\ & & & 1 & -1 & | & 0 & \cdots & 0 \end{bmatrix}$$

と取ると，(2.30) は検定問題 (2.15) の形になる．

次の記号を用いる：

$$\bar{Y}_{i\cdot} = \sum_j Y_{ij}/b, \quad \bar{Y}_{\cdot j} = \sum_i Y_{ij}/a, \quad \bar{Y}_{\cdot\cdot} = \sum_{ij} Y_{ij}/ab = \sum_i \bar{Y}_{i\cdot}/a = \sum_j \bar{Y}_{\cdot j}/b.$$

分散不等式 $\sum_k (x_k - a)^2 \geq \sum_k (x_k - \bar{x})^2$ を用いて，

$$\sum_{ij} (Y_{ij} - \alpha_i - \beta_j)^2 \geq \sum_j \sum_i \{(Y_{ij} - \alpha_i) - (\bar{Y}_{\cdot j} - \bar{\alpha}_\cdot)\}^2$$

$$= \sum_i \sum_j \{(Y_{ij} - \bar{Y}_{\cdot j}) - (\alpha_i - \bar{\alpha}_{\cdot})\}^2$$
$$\geq \sum_{ij} (Y_{ij} - \bar{Y}_{\cdot j} - \bar{Y}_{i\cdot} + \bar{Y}_{\cdot\cdot})^2$$

を得る．この 2 つの不等式の等号は

$$\beta_j = \bar{Y}_{\cdot j} - \bar{\alpha}_{\cdot}, \qquad \alpha_i - \bar{\alpha}_{\cdot} = \bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}$$

ならば成立するが，たとえば

$$\alpha_i = \bar{Y}_{i\cdot}, \qquad \beta_j = \bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}$$

とすればよい．このような解は一意ではないことに注意せよ．結局 R_0^2 は

$$R_0^2 = \min_{\alpha_i, \beta_j} \sum_{ij} (Y_{ij} - \alpha_i - \beta_j)^2 = \sum_{ij} (Y_{ij} - \bar{Y}_{\cdot j} - \bar{Y}_{i\cdot} + \bar{Y}_{\cdot\cdot})^2 =: S_E$$

である．同様に，

$$R_1^2 = \min_{\alpha_1, \beta_j} \sum_{ij} (Y_{ij} - \alpha_1 - \beta_j)^2 = \sum_{ij} (Y_{ij} - \bar{Y}_{\cdot j})^2$$

である．したがって，容易にわかるように，

$$R_1^2 - R_0^2 = b \sum_i (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 =: S_A$$

である．S_E は誤差変動，S_A は行間変動と呼ばれる．

問 6. rank $X = a + b - 1$ であることを示せ．[基本変形によって，X は行列

$$\tilde{X} = \left[\begin{array}{cccc|c} \mathbf{u}_b & & & O & K_b \\ & \mathbf{u}_b & & & K_b \\ & & \ddots & & \vdots \\ & O & & \mathbf{u}_b & K_b \end{array}\right]$$

に変形される．ここで，

$$\mathbf{u}_b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbf{R}^b, \quad K_b = \begin{bmatrix} 1 & | & 0 & \cdots & 0 \\ -- & + & -- & --- & -- \\ -1 & | & & & \\ \vdots & | & & I_{b-1} & \\ -1 & | & & & \end{bmatrix}$$

である．明らかに $\tilde{X}$ の第 $a+1$ 列以外の $a+b-1$ 列は線形独立である．また，第 $a+1$ 列はそれ以外の列の線形結合で書けるから $\tilde{X}$ の階数は $a+b-1$ である．]

rank $H = a-1$ だから，定理 18 をもとに，

$$F' = \frac{S_A/(a-1)}{S_E/\{(a-1)(b-1)\}} \geq f'_\alpha$$

のとき仮説 H'_0 を棄却する．ここで，f'_α は F 分布 $F(a-1,(a-1)(b-1))$ の上側 $100\alpha\%$ 点である．

同様に $B_1, ..., B_b$ の効果に差があるかどうかの検定は

$$H''_0: \ \beta_1 = \cdots = \beta_b \text{ vs. } H''_1: \ \text{ある } j_1, j_2 \text{に対して} \beta_{j_1} \neq \beta_{j_2}$$

と表され，

$$F'' = \frac{S_B/(b-1)}{S_E/\{(a-1)(b-1)\}} \geq f''_\alpha$$

のとき仮説 H''_0 は棄却される．ここで，

$$S_B = a\sum_j (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2$$

で列間変動と呼ばれる．また f''_α は F 分布 $F(b-1,(a-1)(b-1))$ の上側 $100\alpha\%$ 点である．

検定統計量 F', F'' を求める手順は分散分析表にまとめる．全変動 $S_T := \sum_{ij}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2$ の分解

$$S_T = S_E + S_A + S_B$$

が成り立つので，変動のひとつは他の変動から求めることができる．

	平方和	自由度	平均平方和	F
行間	S_A	$a-1$	$S_A/(a-1)$	F'
列間	S_B	$b-1$	$S_B/(b-1)$	F''
誤差	S_E	$(a-1)(b-1)$	$S_E/\{(a-1)(b-1)\}$	
全変動	S_T	$ab-1$		

問 7. 3種類の薬品 A_i $(i=1,2,3)$ を5種類の植物 B_j $(j=1,...,5)$ に一定期間与え，成育の度合い x_{ij} を調べたところ，つぎのデータが得られた：$x_{11}=10, x_{12}=16, x_{13}=16, x_{14}=18, x_{15}=16, x_{21}=15, x_{22}=14, x_{23}=14, x_{24}=13, x_{25}=12, x_{31}=20, x_{32}=18, x_{33}=19, x_{34}=17, x_{35}=18$. 薬品の効果による差があるか，有意水準 5%で検定せよ．また植物による差があるか検定せよ．[$F'=5.48$, $f'_{0.05}=4.46$ (p 値は 0.032) ゆえ有意で，薬品による差が認められる．$F''=0.165$, $f''_{0.05}=3.84$ (p 値は 0.950) となり，植物による差はないといえる．]

文献

1) Adams, R.A.: Sobolev spaces. New York: Academic Press 1975
2) Amari, Shun-ichi: Differential-geometrical methods in statistics. Lecture Notes in Statistics, 28. Springer-Verlag, New York, 1985.
3) Amari, S.-I.; Barndorff-Nielsen, O. E.; Kass, R. E.; Lauritzen, S. L.; Rao, C. R. : Differential geometry in statistical inference. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 10. Institute of Mathematical Statistics, Hayward, CA, 1987.
4) Bhattacharya, R. N., Ghosh, J. K.: On the validity of the formal Edgeworth expansion. Ann. Statist. 6 (1978), no. 2, 434–451. Correction: Ann. Statist. 8 (1980), no. 6, 1399.
5) Billingsley, P.: Convergence of probability measures. John Wiley & Sons, Inc., New York-London-Sydney 1968
6) Dharmadhikari, S. W., Fabian, V., Jogdeo, K.: Bounds on the moments of martingales. Ann. Math. Statist. **39**, 1719–1723 (1968)

7) Ferguson, Thomas S.: Mathematical statistics: A decision theoretic approach. Probability and Mathematical Statistics, Vol. 1 Academic Press, New York-London 1967
8) Ferguson, Th.S.: A course in large sample theory. London Weinheim New York Tokyo Melbourne Madras: Chapman & Hall 1996
9) 舟木直久: 確率論. 朝倉書店 2004
10) 早川毅: 回帰分析の基礎 (統計ライブラリー) 朝倉書店 1986.
11) Huber, P.J.: The behavior of maximum likelihood estimates under nonstandard conditions. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability **1**, 221-233 (1967) Berkeley: University of California Press
12) 稲垣宣生: 数理統計学. 改訂版 裳華房 2003.
13) 伊藤 清: 確率論. 岩波基礎数学選書岩波書店 1991.
14) obayashi, S., Nomizu, K.: Foundations of differential geometry. Interscience tracts in pure and applied mathematics 15. Interscience Publishers 1963.
15) Konishi, S., Kitagawa, G.: Generalised information criteria in model selection. Biometrika 83 (1996), no. 4, 875–890
16) 工藤弘吉: 数理統計学. 現代数学講座 6-C, 共立出版 1957
17) 草間時武: 統計学. サイエンスライブラリ理工系の数学 6. サイエンス社 1975.
18) Lehmann, E. L.: Testing statistical hypotheses. Second edition. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. (渋谷政昭，竹内　啓　訳：統計的検定論. (原著初版)：岩波 1969)
19) Lehmann, E. L.; Casella, George: Theory of point estimation. Second edition. Springer Texts in Statistics. Springer-Verlag, New York, 1998.
20) Lehmann, E.L.: Elements of large-sample theory. New York Berlin Heidelberg: Springer 1999
21) 前園宜彦: 統計的推測の漸近理論. 九州大学出版会 2001
22) 溝畑　茂: ルベーグ積分. 岩波全書，岩波書店 1980
23) 鍋谷清治: 数理統計学. 共立講座現代の数学 33. 共立出版 1978.
24) 西尾真喜子: 確率論. 実教出版 1978.
25) 野水克己: 現代微分幾何入門. 基礎数学選書 25. 裳華房 1981.
26) Pfanzagl, J.: Parametric statistical theory. Berlin New York: Walter de Gruyter 1994
27) Parthasarathy, K. R.: Probability measures on metric spaces. Academic Press, Inc., New York-London 1967
28) Rao, C.R.: Linear statistical inference and its applications. 2nd ed. Wiley 1973 (奥野忠一 他訳. 統計的推測とその応用 （原著第 2 版）: 東京図書 1977)
29) Schweizer, B., Sklar, A.: Probabilistic metric spaces. North-Holland Publishing Co., New York, 1983.
30) Serfling, R.S.: Approximation theorems of mathematical statistics. New York: Wiley 1980
31) 柴田義貞: 正規分布–特性と応用. 東京大学出版会 1981

32) 白旗慎吾: 統計解析入門. 共立出版 1992.
33) 鈴木 武, 山田作太郎：数理統計学：基礎から学ぶデータ解析 第 2 版内田老鶴圃 1998
34) 高松俊朗: 数理統計学入門. 学術図書出版社 1977
35) 竹内　啓: 数理統計学. 東洋経済新報社 1963
36) 竹内啓 他編：統計学辞典. 東洋経済新報社 1989.
37) 竹村彰通：現代数理統計学. 創文社現代経済選書 8. 創文社 1991
38) 東京大学教養学部統計学教室編: 自然科学の統計学. 東京大学出版会 1992.
39) van der Vaart, A.W.: Asymptotic statistics. Cambridge University Press 1998
40) 柳川尭ノンパラメトリック法. 新統計学シリーズ 9. 培風館 1982.
41) 柳井晴夫, 竹内 啓: 射影行列・一般逆行列・特異値分解. UP 応用数学選書 10. 東京大学出版会 1983.

ASAKURA.TEX(h150910)