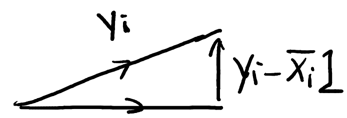
A multivariate observation: (p variables, n observations)

$$X_{(n,p)} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} = [y_1 \,\vdots\, y_2 \,\vdots\, \cdots \,\vdots\, y_p]$$

Geometric explanation: $1_n' = [1, 1, \cdots, 1]$

$$y_i'\left(\frac{1}{\sqrt{n}}1\right)\frac{1}{\sqrt{n}}1 = \frac{x_{1i}+\cdots+x_{ni}}{n}1 = \bar{x}_i 1$$



$$d_i = y_i - \bar{x}_i 1 = \begin{pmatrix} x_{1i} - \bar{x} \\ \vdots \\ x_{ni} - \bar{x} \end{pmatrix}$$ (perpendicular, deviation 1)

$$d_i' d_k = n s_{ik}$$

Result 1: $E(\bar{X}) = \mu$, $Cov(\bar{X}) = \frac{1}{n}\Sigma$

PROOF: $E(\bar{X}) = \frac{1}{n}E(X_1 + \cdots + X_n) = \frac{1}{n}\cdot n \cdot \mu = \mu$

$$(\bar{X}-\mu)(\bar{X}-\mu)' = \frac{1}{n^2}\sum_{i=1}^{n}(X_i-\mu)\sum_{j=1}^{n}(X_j-\mu)'$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i-\mu)(X_j-\mu)'$$

$$Cov(\bar{X}) = E(\bar{X}-\mu)(\bar{X}-\mu)' = \frac{1}{n^2}\left(\sum_{i=1}\sum_{j=1}(X_i-\mu)(X_j-\mu)'\right)$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}E(X_j-\mu)(X_j-\mu)'\right) = \frac{1}{n^2}\cdot n\Sigma = \frac{1}{n}\Sigma$$

Result 4.4: $X = \begin{pmatrix} X_1 \\ \hdashline X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \hdashline \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$

$$\Rightarrow X_1 \sim N(\mu_1, \Sigma_{11})$$

Result 4.5 X_1, X_2 are independent iff $\Sigma_{12} = 0$

Exercise 4.14

Result 4.6 Conditional $X_2 = x_2$:

$$\exists A,\ \text{s.t. } A\Sigma A' = \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$$

$$A = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \quad \therefore A\mu = \begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}$$

$$\therefore AX = \begin{pmatrix} X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \\ \hdashline X_2 \end{pmatrix} \text{ and } X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \perp\!\!\!\perp X_2$$

given $X_2 = x_2$,

we have $X_1 - \Sigma_{12}\Sigma_{22}^{-1}x_2 \sim N(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2.\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$

Result 4.7: $(X-\mu)\Sigma^{-1}(X-\mu)$ is distributed as χ_p^2 with p freedom

why?

$\Sigma^{-1/2}(X-\mu) \sim N(0, I)$

χ_p^2

Result 4.8: $V_1 = \sum b_i X_i,\ V_2 = \sum c_i X_i$

$$\therefore \begin{pmatrix} V_1 \\ \hdashline V_2 \end{pmatrix} = \begin{pmatrix} b_1 & \cdots & b_n \\ c_1 & \cdots & c_n \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\therefore \begin{pmatrix} b \\ \hdashline c \end{pmatrix} \begin{pmatrix} \Sigma_1 & & \\ & \ddots & \\ & & \Sigma_n \end{pmatrix} (b \,\vdots\, c) = \begin{pmatrix} b'b\Sigma & b'c\Sigma \\ c'b\Sigma & c'c\Sigma \end{pmatrix}$$

Joint Density:

$$p(X) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{n/2}} e^{-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)}$$

$$= \frac{1}{(2\pi)^{n/2}|\Sigma|^{n/2}} e^{-\frac{1}{2}tr(\Sigma^{-1}(X-\mu)(X-\mu)')}$$

$$tr(\Sigma^{-1}(X-\mu)(X-\mu)') = \sum_{i=1}^{n} tr(\Sigma^{-1}(x_i-\mu)(x_i-\mu)')$$

$$= \sum_{i=1}^{n} tr\left(\Sigma^{-1}(x_i-\bar{x}+\bar{x}-\mu)(x_i-\bar{x}+\bar{x}-\mu)'\right)$$

$$\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu)(x_i-\bar{x}+\bar{x}-\mu)' = \sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})' + n(\bar{x}-\mu)(\bar{x}-\mu)'$$

$$+ 2\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})(\bar{x}-\mu)'}_{=0}$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})' + n(\bar{x}-\mu)(\bar{x}-\mu)'$$

Likelihood function:

$$L(\mu,\Sigma) = \frac{1}{(2\pi)^{n/2}\cdot|\Sigma|^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n} tr(\Sigma^{-1}(x_i-\bar{x})(x_i-\bar{x})') - \frac{n}{2}(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)}$$

1° $\hat{\mu} = \bar{x}$ is obvious

2° for $\hat{\Sigma}$: $L(\hat{\mu},\Sigma) = \frac{1}{(2\pi)^{n/2}\cdot|\Sigma|^{n/2}} e^{-\frac{1}{2}tr\left(\Sigma^{-1}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})'\right)}$

Result 4.10: $\frac{1}{|\Sigma|^b} e^{-tr(\Sigma^{-1}B)/2} \le \frac{1}{|B|^b}(2b)^{pb} e^{-pb}$

PROOF: $tr(\Sigma^{-1}B) = tr(B^{1/2}\Sigma^{-1}B^{1/2}) = \sum_{i=1}^{n} \eta_i$

$|\Sigma^{-1}B| = \prod_{i=1}^{n} \eta_i \qquad \therefore \frac{1}{|\Sigma|} = \frac{\prod_{i=1}^{n} \eta_i}{|B|}$

$\therefore \frac{1}{|\Sigma|^b} e^{-tr(\Sigma^{-1}B)/2} = \frac{\prod_{i=1}^{n} \eta_i^b}{|B|^b} e^{-\frac{1}{2}\sum_{i=1}^{n}\eta_i}$

for each $\eta_i^b e^{-\eta_i/2}$, we have $\frac{\partial}{\partial \eta_i}(b\ln\eta_i - \eta_i/2) = 0$

and $\eta_i^b e^{-\eta_i/2} \le (2b)^b e^{-b}$

$\therefore \frac{1}{|\Sigma|^b} e^{-\sum_{i=1}^{p}\eta_i} \le \frac{1}{|B|^b}(2b)^{pb} e^{-pb}$

Upper Bound is reached if $\Sigma = (1/2b)B$

Result 4.11: $\hat{\Sigma} = (1/n)\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})'$

Wishart: $X = [X_1, \cdots, X_n]$ and $M = XX' = \sum_{i=1}^{n} X_i X_i' \sim W_p(n, \Sigma)$

property: (1) $B'MB_{(p\times m)} \sim W_m(n, B'\Sigma B)$

(2) $M_1 + \cdots + M_K \sim W_p(n_1 + \cdots + n_K, \Sigma)$

(3) $E(M_n) = n\Sigma$

Corollary: $\frac{\alpha' M \alpha}{\alpha' \Sigma \alpha} \sim \chi_n^2, \quad \frac{\alpha'\Sigma^{-1}\alpha}{\alpha' M^{-1}\alpha} \sim \chi^2_{n-p+1}$

With Large n: $\sqrt{n}(\bar{X} - \mu) \sim N_p(0, \Sigma)$

$n(\bar{X} - \mu)' S^{-1}(\bar{X} - \mu) \sim \chi_p^2$

4.2

(a): $\mu = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ $\Sigma = \begin{pmatrix} 2 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{pmatrix}$, $p(X) = \frac{1}{(2\pi)|\Sigma|} e^{-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)}$

(b) $\Sigma^{-1} = \frac{2}{3}\begin{pmatrix} 1 & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{\sqrt{2}}{3} \\ -\frac{\sqrt{2}}{3} & \frac{4}{3} \end{pmatrix}$

$\therefore (X-\mu)'\Sigma^{-1}(X-\mu) = \frac{2}{3}x_1^2 - \frac{2\sqrt{2}}{3}x_1(x_2-2) + \frac{4}{3}(x_2-2)^2$

(c) $\alpha_{50,\chi_2^2} = 1.38$,

Contour: $\frac{2}{3}x_1^2 - \frac{2\sqrt{2}}{3}x_1(x_2-2) + \frac{4}{3}(x_2-2)^2 \le 1.38$

4.5:

(a) Apply the universal case

$X_1 | X_2 = x_2 \sim N(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mu_2 - x_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12})$

$\Rightarrow$: $X_1 | X_2 = x_2 \sim N(\frac{\sqrt{2}}{2}(x_2-2), \frac{3}{2})$

(b) $\tilde{\Sigma} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 5 \end{pmatrix}$ $\quad \Sigma_1^{-1} = \begin{pmatrix} 1 & \\ & \frac{1}{2} \end{pmatrix}$

$X_2 | X_1 = x_1, X_3 = x_3 \sim N(1 + (x_1+3) + \frac{1}{2}(x_2-4), 1)$

(c) $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$ $\quad \Sigma_1^{-1} = \frac{1}{2}\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$

$\therefore \mu_3 - \Sigma_{12}\Sigma_1^{-1}(\mu - x) = 1 + \frac{1}{2}(x_1-2) + \frac{1}{2}(x_2+3)$

$\Sigma_3 - \Sigma_{12}\Sigma_1^{-1}\Sigma_{12} = 2 - \frac{3}{2} = \frac{1}{2}$

$\therefore X_3 | X_1 = x_1, X_2 = x_2 \sim N(1 + \frac{1}{2}(x_1-2) + \frac{1}{2}(x_2+3), \frac{1}{2})$

4.16 (a) $V_1 \sim N(0, \frac{1}{4}\Sigma)$

$V_2 \sim N(0, \frac{1}{4}\Sigma)$

(b) $Cov(V_1, V_2) = Cov(\frac{1}{4}X_1 - \frac{1}{4}X_2 + \frac{1}{4}X_3 - \frac{1}{4}X_4, \frac{1}{4}X_1 + \frac{1}{4}X_2 - \frac{1}{4}X_3 - \frac{1}{4}X_4)$

$= \frac{1}{16}Var(X_1) - \frac{1}{16}Var(X_2) - \frac{1}{16}Var(X_3) + \frac{1}{16}Var(X_4)$

$= 0$

$\therefore \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \sim N\left(0, \begin{pmatrix} \Sigma & \\ & \Sigma \end{pmatrix}\right)$

4.21 (a) $\bar{X} \sim N(\mu, \frac{1}{60}\Sigma)$

(b) $\Sigma^{-1/2}(X_1 - \mu) \sim N(0, I)$, $\therefore (X_1-\mu)'\Sigma^{-1}(X_1-\mu) \sim \chi^2_{60}$

(c) $\Sigma^{-1/2}(\bar{X} - \mu) \sim N(0, \frac{1}{60}I)$, $\therefore 60(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu) \sim \chi^2_{60}$

(d) Approximately, $S^{-1} \xrightarrow{P} \Sigma^{-1}$

With Slutsky Theorem, $n(\bar{X}-\mu) S^{-1}(\bar{X}-\mu) \sim \chi^2_{60}$

Regression Models

- **Classical Model**

  $$\underset{(n\times1)}{Y} = \underset{n\times(r+1)}{Z}\ \underset{(r+1)\times1}{\beta} + \underset{n\times1}{\varepsilon}$$

  $$E(\varepsilon) = \underset{(n\times1)}{0}, \quad Cov(\varepsilon) = \sigma^2 \underset{(n\times n)}{I}$$

  β, σ^2 unknown parameter, Z design matrix

- Result 7.1 Least squares estimate of β is given by $\hat{\beta} = (Z'Z)^{-1}Z'y$

  $\hat{y} = Z\hat{\beta} = Hy$, where $H = Z(Z'Z)^{-1}Z'$.

  Residuals $\hat{\varepsilon} = y - \hat{y} = (I-H)y$ satisfies $Z'\hat{\varepsilon} = 0$ and $\hat{y}\cdot\hat{\varepsilon} = 0$

  $$RSS = \sum_{i=1}^{n}(y_i - \hat{\alpha} - \hat{\beta}x_i)^2 = \hat{\varepsilon}'\hat{\varepsilon} = y'[I - Z(Z'Z)^{-1}Z']y = y'y - y'Z\hat{\beta}$$

- Sum of square decomposition

  $1°\ y'y = (\hat{y} + y - \hat{y})'(\hat{y} + y - \hat{y}) = (\hat{y} + \hat{\varepsilon})'(\hat{y} + \hat{\varepsilon}) = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}$

  $2°\ 0 = 1'\hat{\varepsilon} = \sum_{i=1}^{n}\hat{\varepsilon}_i = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\hat{y}_i \Rightarrow \bar{y}_i = \bar{\hat{y}}_i$

  $$\overset{1°+2°}{\Longrightarrow}\ y'y - n(\bar{y})^2 = \hat{y}'\hat{y} - n(\bar{\hat{y}})^2 + \hat{\varepsilon}'\hat{\varepsilon}$$

  or $$\sum_{j=1}^{n}(y_j - \bar{y})^2 = \sum_{j=1}^{n}(\hat{y}_j - \bar{y})^2 + \sum_{j=1}^{n}\hat{\varepsilon}_j^2$$

  $$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

  $$TSS = RSS + ESS$$

  $$R^2 = 1 - \frac{ESS}{TSS} = \frac{RSS}{TSS}$$

- Result 7.2:

1° $E(\hat{\beta})=\beta$ and $Cov(\hat{\beta})=\sigma^2(Z'Z)^{-1}$

2° $E(\hat{\varepsilon})=0$ and $Cov(\hat{\varepsilon})=\sigma^2[I-Z(Z'Z)^{-1}Z']=\sigma^2(I-H)$

3° $E(\hat{\varepsilon}'\hat{\varepsilon})=(n-r-1)\sigma^2$

PROOF:

Lemma: $E(Y'AY)=E((Y-E(Y)+E(Y))'A(Y-E(Y)+E(Y)))$

$$=E((Y-E(Y))A(Y-E(Y)))+E(Y)'AE(Y)$$

$$=tr(AE(Y-E(Y))(Y-E(Y))')+E(Y)'AE(Y)$$

$$=tr(A\,Var(Y))+E(Y)'AE(Y)$$

$$\therefore E(\hat{\varepsilon}'\hat{\varepsilon})=E(Y'(I-Z(Z'Z)^{-1}Z')Y)=tr(I-Z(Z'Z)^{-1}Z')\cdot\sigma^2+\beta'Z'[I-Z(Z'Z)^{-1}Z']Z\beta$$

$$=\sigma^2tr(I)-\sigma^2tr((Z'Z)^{-1}(Z'Z))=(n-(r+1))\sigma^2$$

4° $\hat{\beta}$ and $\hat{\varepsilon}$ are uncorrelated

PROOF: $Cov(\hat{\varepsilon},\hat{\beta})=Cov((I-H)y,Hy)=0$

· Result 7.3: (Gauss Markov theorem) $c'\hat{\beta} = c_0\hat{\beta}_0 + c_1\hat{\beta}_1 + \cdots + c_r\hat{\beta}_r$ is the smallest possible variance among all linear estimators of $a'Y = a_1Y_1 + \cdots + a_nY_n$

PROOF $c'\beta = c_0\beta_0 + \cdots + c_r\beta_r$

Estimator: $c'\hat{\beta} = c_0\hat{\beta}_0 + c_1\hat{\beta}_1 + \cdots + c_r\hat{\beta}_r = \underbrace{c'(Z'Z)^{-1}Z'}_{a'}Y$

$E(c'\hat{\beta}) = c'\beta$, $V(c'\hat{\beta}) = a'\sigma^2 I a = \sigma^2 a'a$

Alternative Estimator: $d'Y$ with $E(d'Y) = d'Z\beta = c'\beta \quad \forall\beta \Rightarrow d'Z = c'$

$$V(d'Y) = d'\sigma^2 I d = \sigma^2 d'd = \sigma^2[(a+d-a)'(a+d-a)]$$

$$= \sigma^2[a'a + (d-a)(d-a) + 2a'(d-a)] = \sigma^2[a'a + (d-a)'(d-a)]$$

$$\geq V(c'\hat{\beta})$$

$$a'(d-a) = c'(Z'Z)^{-1}Z'(Z^{-1} - Z(Z'Z)^{-1})c$$

$$= c'(Z'Z)^{-1}c - c'(Z'Z)^{-1}c = 0$$

$\Longrightarrow c'\hat{\beta}$ is the BLUE of $c'\beta$

Inference about a mean vector

$$H_0: \mu = \mu_0 \quad \text{v.s.} \quad H_1: \mu \neq \mu_0$$

$$T = \frac{(\bar{X} - \mu)}{s/\sqrt{n}} \quad \text{where} \quad \bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \quad \text{and} \quad s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2$$

$$\{\text{Do not reject } H_0: \mu = \mu_0 \text{ at level } \alpha\} \Leftrightarrow \left\{\left|\frac{\bar{X} - \mu_0}{s/\sqrt{n}}\right| \leq t_{n-1}^2(\alpha/2)\right\}$$

$$\text{or} \quad \bar{X} - t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}$$

Multivariate Version:

$$H_0: \mu = \mu_0 \quad \text{v.s.} \quad H_1: \mu \neq \mu_0$$

$$T^2 = (\bar{X} - \mu_0)'\left(\frac{1}{n}S\right)^{-1}(\bar{X} - \mu_0) \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

$$\text{where} \quad \bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \quad \text{and} \quad S = \frac{1}{n-1}\sum_{i=1}^{n}(X_j - \bar{X})(X_j - \bar{X})'$$

Hotelling's T^2

$$T^2_{p,n-1} \sim N_p(0,\Sigma)'\left[\frac{1}{n-1} W_{p,n-1}(\Sigma)\right]^{-1} N_p(0,\Sigma)$$

analogous to

$$t_{n-1}^2 = \sqrt{n}(\bar{X} - \mu_0)(s^2)^{-1}\sqrt{n}(\bar{X} - \mu_0)$$

· Affine Invariance

$$Y = CX + d \Rightarrow \bar{y} = C\bar{x} + d, \quad S_y = \frac{1}{n-1}\sum_{i=1}^{n}(y_j - \bar{y})(y_j - \bar{y})' = CSC'$$

$$\mu_Y = E(Y) = C\mu + d$$

$$\therefore T^2 = n(C(\bar{x}-\mu))' \cdot (CSC')^{-1}(C(\bar{x}-\mu)) = n(\bar{x}-\mu)S^{-1}(\bar{x}-\mu)$$

Likelihood Ratio Test ($H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$)

$$L(\mu_0, \Sigma) = \frac{1}{(2\pi)^{np/2}(\Sigma)^{n/2}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n}(x_j - \mu_0)'\Sigma^{-1}(x_j - \mu_0)\right)$$

$$\Rightarrow \text{Likelihood Ratio} = \Lambda = \frac{\max_{\Sigma} L(\mu_0, \Sigma)}{\max_{\mu, \Sigma} L(\mu, \Sigma)} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{n/2}$$

As result 4.10 proved,
$$\begin{cases} \hat{\mu} = \sum_{i=1}^{n} \bar{x} & \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})' \\ \mu_0 = \mu_0, & \hat{\Sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)(x_i - \mu_0)' \end{cases}$$

Reject $H_0: \mu = \mu_0$ if $\Lambda = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{n/2} = \left(\frac{\left|\sum_{j=1}^{n}(x_j - \bar{x})(x_j - \bar{x})'\right|}{\left|\sum_{j=1}^{n}(x_j - \mu_0)(x_j - \mu_0)'\right|}\right)^{n/2} < C_\alpha$

Result 5.1: $\Lambda^{2/n} = \left(1 + \dfrac{T^2}{(n-1)}\right)^{-1}$

PROOF Let $(p+1)\times(p+1)$ matrix

$$A = \left(\begin{array}{c|c} \sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})' & \sqrt{n}(\bar{x}-\mu_0) \\ \hline \sqrt{n}(\bar{x}-\mu_0)' & -1 \end{array}\right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

$$|A| = |A_{22}|\,|A_{11} - A_{12}A_{22}^{-1}A_{21}| = |A_{11}|\,|A_{22} - A_{21}A_{11}^{-1}A_{12}|$$

$$\Rightarrow (-1)\left|\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})' + n(\bar{x}-\mu_0)(\bar{x}-\mu_0)'\right| = \left|\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'\right| \cdot \left|-1 - n(\bar{x}-\mu_0)'\left(\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'\right)^{-1}(\bar{x}-\mu_0)\right|$$

$$\Downarrow$$

$$\sum_{j=1}^{n}(x_j-\mu_0)(x_j-\mu_0)' = \sum_{j=1}^{n}(x_j-\bar{x}+\bar{x}-\mu_0)(x_j-\bar{x}+\bar{x}-\mu_0)'$$

$$= \sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})' + n(\bar{x}-\mu_0)(\bar{x}-\mu_0)'$$

$$\Downarrow$$

$$(-1)\left|1 + \frac{T^2}{(n-1)}\right|$$

$$\therefore (-1)\Big|\underbrace{\sum_{j=1}^{n}(x_j-\mu_0)(x_j-\mu_0)'}_{n\hat{\Sigma}_0}\Big| = \Big|\underbrace{\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'}_{n\hat{\Sigma}}\Big|(-1)\left(1+\frac{T^2}{(n-1)}\right)$$

$$\Rightarrow \Lambda^{2/n} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1+\frac{T^2}{(n-1)}\right)^{-1}$$

#

Result 5.2 When n is large, $-2\ln\Lambda = -2\ln\left(\dfrac{\max\limits_{\theta\in\Theta_0} L(\theta)}{\max\limits_{\theta\in\Theta} L(\theta)}\right)$ is approximately $\chi^2_{\nu-\nu_0}$

Here $\nu - \nu_0 = (\dim\Theta) - (\dim\Theta_0)$

Inference about Regression Model

- Result 7.4: $Y = Z\beta + \varepsilon$, where Z has rank $r+1$ and $\varepsilon \sim N_n(0, \sigma^2 I)$

  1° $\hat{\beta} = (Z'Z)^{-1}Z'Y \sim N_{r+1}(\beta, \sigma^2(Z'Z)^{-1})$

  PROOF $\hat{\beta} = \underbrace{(Z'Z)^{-1}Z'Z\beta}_{\beta} + \underbrace{(Z'Z)^{-1}Z'\varepsilon}_{N_{r+1}(0, \sigma^2(Z'Z)^{-1})} \sim N_{r+1}(\beta, \sigma^2(Z'Z)^{-1})$

  2°. $n\hat{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon} \sim \sigma^2 \chi^2_{n-r-1}$

  PROOF $\hat{\varepsilon}'\hat{\varepsilon} = (y-\hat{y})'(y-\hat{y}) = y'(I-H)y$

  $\because$ $I-H$ is idempotent, it can be diagonalized by $U^T\begin{pmatrix}1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0\end{pmatrix}U$

  $tr(I-H) = tr(I) - tr(X(X'X)^{-1}X') = n - tr(I_{r+1}) = n-r-1$

  $\therefore y'(I-H)y = \sum_{i=1}^{n-r-1} \tilde{y}_i^2 \sim \chi^2_{n-r-1}$

- Result 7.5: Let $Y = Z\beta + \varepsilon$, then $100(1-\alpha)$ percent confidence region for β is given by

$$(\beta-\hat{\beta})Z'Z(\beta-\hat{\beta}) \leq (r+1)s^2 F_{r+1,n-r-1}(\alpha)$$

$F_{r+1,n-r-1}(\alpha)$ is the upper (100α)th percentile of $F_{r+1,n-r-1}$

PROOF: Let $V = (Z'Z)^{1/2}(\hat{\beta}-\beta)$, then

$$\begin{cases} E(V) = 0 \\ Cov(V) = (Z'Z)^{1/2} Cov(\hat{\beta})(Z'Z)^{1/2} = \sigma^2 I \end{cases}$$

since V is normally-distributed, $V'V \sim \sigma^2\chi^2_{r+1}$

$\Rightarrow (\beta-\hat{\beta})'Z'Z(\beta-\hat{\beta}) \sim \sigma^2\chi^2_{r+1}$, and $(n-r-1)s^2 = \varepsilon'\varepsilon \sim \sigma^2\chi^2_{n-r-1}$

$\Rightarrow [V'V/(r+1)]/s^2$ is $F_{r+1,n-r-1}$ distribution

$\Rightarrow (\beta-\hat{\beta})'Z'Z(\beta-\hat{\beta}) \leq (r+1)\cdot s^2 F_{r+1,n-r-1}(\alpha)$

Likelihood Ratio Test for β

$$H_0: \beta_{q+1}=\beta_{q+2}=\cdots=\beta_r=0 \quad \text{v.s.} \quad H_1: \beta_i \neq 0$$

- $\underset{n\times(q+1)}{Z_1}$, $Z=[Z_1 \mid Z_2]$, $\underset{n\times(r-q)}{Z_2}$, $\beta=\begin{pmatrix} \underset{(q+1)\times 1}{\beta_{(1)}} \\ \hline \underset{(r-q)\times 1}{\beta_{(2)}} \end{pmatrix}$. $Y=Z\beta+\varepsilon=Z_1\beta_{(1)}+Z_2\beta_{(2)}+\varepsilon$
- Extra Sum of square $=SS_{res}(Z_1)-SS_{res}(Z)$

$$=(y-Z_1\hat{\beta}_{(1)})'(y-Z_1\hat{\beta}_{(1)})-(y-Z\hat{\beta})'(y-Z\hat{\beta})$$

$$\text{F-stats}=\frac{(SS_{res}(Z_1)-SS_{res}(Z))/(r-q)}{SS_{res}(Z)/(n-r-1)}\sim F_{r-q,n-r-1}$$

- Result 7.6: The likelihood ratio test of $H_0: \beta_{(2)}=0$ is equivalent to a test of H_0 based on F-statistics. We reject H_0 if

$$\frac{(SS_{res}(Z_1)-SS_{res}(Z))/(r-q)}{SS_{res}(Z)/(n-r-1)} > F_{r-q,n-r-1}(\alpha)$$

$F_{r-q,n-r-1}(\alpha)$ is the upper (100α)th percentile of $F_{r-q,n-r-1}$

PROOF: $L(\beta,\sigma^2)=\frac{1}{(2\pi)^{n/2}\sigma^n}\exp(-(y-Z\beta)'(y-Z\beta)/2\sigma^2)\leq\frac{1}{(2\pi)^{n/2}\hat{\sigma}^n}e^{-n/2}$

With $\hat{\beta}=(Z'Z^{-1})Z'y$, $\hat{\sigma}^2=(y-Z\hat{\beta})'(y-Z\hat{\beta})/n$

$$\max_{\beta_{(1)},\sigma^2} L(\beta_{(1)},\sigma^2)=\frac{1}{(2\pi)^{n/2}\hat{\sigma}_1^n}e^{-n/2}$$

with $\hat{\beta}_{(1)}=(Z_1'Z_1)^{-1}Z_1'y$, $\hat{\sigma}_1^2=(y-Z_1\hat{\beta}_{(1)})'(y-Z_1\hat{\beta}_{(1)})/n$

Likelihood Ratio: $\frac{\max L(\beta_{(1)},\sigma^2)}{\max L(\beta,\sigma^2)}=\left(\frac{\hat{\sigma}_1}{\hat{\sigma}}\right)^{-n/2}=\left(1+\frac{\hat{\sigma}_1^2-\hat{\sigma}^2}{\hat{\sigma}^2}\right)^{-n/2}$

$$T(\hat{\beta},\hat{\sigma})=\frac{n(\hat{\sigma}_1^2-\hat{\sigma}^2)/(r-q)}{n\hat{\sigma}^2/(n-r-1)}=\frac{(SS_{res}(Z_1)-SS_{res}(Z))/(r-q)}{SS_{res}(Z)/(n-r-1)}\sim F_{r-q,n-r-1}$$

Plus: $H_0: C\beta=0$ v.s. $H_1: C\beta\neq 0$ with $C\hat{\beta}\sim N_{r-q}(C\beta,\sigma^2C(Z'Z)^{-1}C')$

$$\frac{(C\hat{\beta})'(C(Z'Z)^{-1}C')^{-1}(C\hat{\beta})}{\sigma^2}>(r-q)F_{r-q,n-r-1}(\alpha)$$

Multivariate Multiple Regression

$$Y_1 = \beta_{01} + \beta_{11} z_1 + \cdots + \beta_{r1} z_r + \varepsilon_1$$
$$Y_2 = \beta_{02} + \beta_{12} z_1 + \cdots + \beta_{r2} z_r + \varepsilon_2$$
$$\vdots \qquad\qquad \vdots$$
$$Y_m = \beta_{0m} + \beta_{1m} z_1 + \cdots + \beta_{rm} z_r + \varepsilon_m$$

Notations:

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} \\ \vdots & \vdots & & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nm} \end{pmatrix} = \left[Y_{(1)} \,\vdots\, Y_{(2)} \,\vdots\, \cdots \,\vdots\, Y_{(m)} \right]$$

$$\beta = \begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0m} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \vdots & \vdots & & \vdots \\ \beta_{r1} & \beta_{r2} & \cdots & \beta_{rm} \end{pmatrix} = \left(\beta_{(1)} \,\vdots\, \beta_{(2)} \,\vdots\, \cdots \,\vdots\, \beta_{(m)} \right)$$

$$\varepsilon = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ \vdots & \vdots & & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nm} \end{pmatrix} = \left[\varepsilon_{(1)} \,\vdots\, \varepsilon_{(2)} \,\vdots\, \cdots \,\vdots\, \varepsilon_{(m)} \right] = \begin{bmatrix} \varepsilon_1' \\ \hline \varepsilon_2' \\ \vdots \\ \hline \varepsilon_n' \end{bmatrix}$$

Model

$$\underset{(n\times m)}{Y} = \underset{n\times(r+1)}{Z}\,\underset{(r+1)\times m}{\beta} + \underset{n\times m}{\varepsilon}, \quad \text{with } E(\varepsilon_{(i)}) = 0 \text{ and } Cov(\varepsilon_{(i)}, \varepsilon_{(k)}) = \sigma_{ik} I$$

m observations on one trial have $\Sigma = \{\sigma_{ik}\}$

different trials are independent

Estimates:

$$\hat{\beta} = \left[\hat{\beta}_{(1)} \,\vdots\, \hat{\beta}_{(2)} \,\vdots\, \cdots \,\vdots\, \hat{\beta}_{(m)} \right] = (Z'Z)^{-1} Z' \left[Y_{(1)} \,\vdots\, Y_{(2)} \,\vdots\, \cdots \,\vdots\, Y_{(m)} \right]$$

or $\hat{\beta} = (Z'Z)^{-1} Z' Y$