Hypothesis Testing

(1) Testing Model:

· Consider $Y = X\beta + \varepsilon$. $\varepsilon \sim N_n(0, \sigma^2 I)$ ⋯ (*)

Def: i) $C(X)$ is estimation space

ii) $C^{\perp}(X)$ is error space.

We're interested in reducing full model (*) to

a reduced model : $Y = X_0 \gamma_0 + \varepsilon$. $C(X_0) \subset C(X)$.

e.g. For $Y_{ij} = \mu + a_i + \varepsilon_{ij}$, under H_0: $a_1 = a_2 = \cdots = a_n$

reduce it to $Y_{ij} = \mu + \varepsilon_{ij}$.

⇒ Express in $E(Y)$:

H_0: $E(Y) \in C(X_0)$ v.s. H_1: $E(Y) \in C(X) \cap C^{c}(X_0)$

Rmk: i) If the reduced model is true. Then:

(*) is true as well.

ii) Under H_0: UMVUE of $E(Y)$ is $M_0 Y$

Note that : For $M = X(X^T X)^- X^T$. $M_0 = X_0 (X_0^T X_0)^- X_0^T$.

If H_0 is true. Then $(M - M_0)Y$ will be small.

⇒ So "whether H_0 is true" depends on the size of $(M-M_0)Y$. i.e. $Y^T(M-M_0)Y = \|(M-M_0)Y\|^2$.

Since Y is random. Consider:

$E\left(\frac{Y^T(M-M_0)Y}{r(M-M_0)}\right)$ as $\|(M-M_0)Y\|^2$'s estimation.

i) If H_0 is true:

$$E\left(\frac{Y^T(M-M_0)Y}{r(M-M_0)}\right) = \sigma^2 + \frac{(X\beta)^T(M-M_0)X\beta}{r(M-M_0)}$$
$$= \sigma^2$$

ii) If H_0 isn't true:

$$E\left(\frac{Y^T(M-M_0)Y}{r(M-M_0)}\right) = \sigma^2 + \frac{(X\beta)^T(M-M_0)X\beta}{r(M-M_0)}$$

Rmk: i) H_0 is true $\Rightarrow$ $\frac{Y^T(M-M_0)Y}{r(M-M_0)\sigma^2} = 1$.

Replace σ^2 by its estimator $\hat{\sigma}^2 =$

$Y^T(I-M)Y / r(I-M)$

ii) $\frac{\|(M-M_0)X\beta\|^2}{\sigma^2} = \frac{\|(I-M_0)X\beta\|^2}{\sigma^2}$ is

crucial for the testing.

Thm. Under the models above:

$$F = \frac{\|(M-M_0)Y\|^2 / r(M-M_0)}{\|(I-M)Y\|^2 / r(I-M)} \sim F\left(r(M-M_0),\ r(I-M),\ \frac{\|(M-M_0)X\beta\|^2}{\sigma^2}\right)$$

If H_0 is true $\Rightarrow$ $\gamma = 0$ (non-central para.)

If H_1 is true $\Rightarrow$ $\gamma \neq 0$.

Then the rejection region is:

$\{ F > F_{1-\alpha}(r(M-M_0), r(I-M)) \}$ (percentile) with level "α".

<u>Pf</u>: 1°) Independence of numerator and denominator is easy to check.

2°) $Y^T(I-M)Y \sim \sigma^2 \chi^2(r(I-M))$ with.

$Y^T(M-M_0)Y \sim \sigma^2 \chi^2(r(M-M_0), \gamma)$

(2) <u>Testing Linear Para. Function</u>:

① Consider : $Y = X\beta + \varepsilon$. $\varepsilon \sim N_n(0, \sigma^2 I)$

We want to test sth. w.r.t the estimable function $\Lambda^T \beta = \rho^T X \beta$:

H_0: $\rho^T X\beta = 0$. v.s. H_1: $\rho^T X\beta \neq 0$.

<u>Rmk</u>: The hypothesis means $X\beta = E(Y) \in N(\rho^T)$ $= C^{\perp}(\rho)$. Restate hypothesis:

H_0: $E(Y) \in C(X) \cap C^{\perp}(\rho)$ v.s H_1: $E(Y) \notin \square$

Find X_0. s.t. $C(X_0) = C(X) \cap C^{\perp}(\rho)$. Then reduce the case to (1). But X_0 not unique.

i) <u>Find X_0</u>:

(a) Find $C(U) = C^{\perp}(\Lambda)$. $\Rightarrow \beta = UY$. $\Rightarrow E(Y) \in C(XU) = C(X_0)$

(b) Set $M = P_{C(X)|C(X)}$. $P = MP + (I-M)P$

$\therefore P^T X\beta = P^T M X\beta$. $E(Y) \perp C(P) \Leftrightarrow E(Y) \perp C(MP)$

restate H_0: $E(Y) \in C(X) \cap C^{\perp}(MP)$.

Thm. $X_0 = (I - M_{MP})X \Rightarrow C(X_0) = C(X) \cap C^{\perp}(MP)$

where $M_{MP} = MP(P^T MP)^- P^T M. = P_{C(MP)|C(MP)}$

Pf: Check: $C((I-M_{MP})X) \subset C(X) \cap C^{\perp}(MP)$

$C((I-M_{MP})X) \supset C(X) \cap C^{\perp}(MP)$

Rmk: It shows one of choice of X_0 is:

$(I - M_{MP}) X \in M^{n\times p}$

Since $C(X) = C(M) \Rightarrow$ Another choice:

$(M - M_{MP}) \in M^{n\times n}$, $M_0 = M - M_{MP}$.

ii) To test hypothesis:

By (1): numerator is $Y^T(M - M_0)Y = Y^T M_{MP} Y$

$$F = \frac{\|M_{MP}Y\|^2 / r(M_{MP})}{\|(I-M)Y\|^2 / r(I-M)} \sim F(r(M_{MP}), r(I-M), \gamma)$$

$$\gamma = \frac{\|M_{MP}X\beta\|^2}{\sigma^2}$$ non-central para. (Under H_0: $\gamma = 0$)

Rmk: To rehaw F: Note that $\widehat{\Lambda^T \beta} = P^T M Y$

(where $\Lambda^T\beta$ is estimable)

Besides. $r(\Delta) = r(M_{MP}) = r(MP)$

<u>Pf:</u> $\Delta^T = P^T X$. Note: $b \in \mathbb{R}^n$. $X^T P b = 0$

$\Leftrightarrow Pb \perp C(X) \Leftrightarrow MPb = 0$

$\therefore C^{\perp}(P^T X) = C^{\perp}(P^T M) \Rightarrow C(P^T X) = C(P^T M)$

$r(\Delta) = r(X^T P) = r(MP)$ $\square$

$$\Rightarrow Y^T M_{MP} Y = \widehat{\beta^T \Delta} (P^T X (X^T X)^- X^T P)^- \widehat{\Delta^T \beta}$$

$$= \widehat{\beta^T \Delta} (\Delta^T (X^T X)^- \Delta)^- \widehat{\Delta^T \beta}$$

$$F = \frac{\widehat{\beta^T \Delta} (\Delta^T (X^T X)^- \Delta)^- \widehat{\Delta^T \beta} / r(\Delta)}{Y^T (I-M) Y / r(I-M)} . \ \gamma = \frac{\beta^T \Delta (\Delta^T (X^T X)^- \Delta) \beta \Delta^T}{\sigma^2}$$

$\sim F(r(\Delta), r(I-M), \gamma)$

<u>Rmk:</u> i) $Cov(\widehat{\Delta^T \beta}) = \sigma^2 \Delta^T (X^T X)^- \Delta$

ii) To test the hypothesis, we need:

a) $\widehat{\Delta^T \beta}$. $Cov(\widehat{\Delta^T \beta})$

b) $r(\Delta)$ c) $\|(I-M)Y\|^2 / r(I-M)$

② <u>Theoretical Complement:</u>

· We now examine $\Delta^T \beta = 0$ When it's not estimable

<u>Thm.</u> If $C(\Delta_0) = C(\Delta) \cap C(X^T)$. $C(U_0) = C^{\perp}(\Delta_0)$

$C(U) = C^{\perp}(\Delta)$. Then: $C(XU_0) = C(XU)$

<u>Rmk:</u> It means $\Delta^T \beta = 0$ or $\Delta_0^T \beta = 0$ induces

the same reduced model. $C(XU) = C(MP)$

Pf: $C(\Lambda_0) \subset C(\Lambda) \Rightarrow C(U) \subset C(U_0)$

$\therefore C(XU) \subset C(XU_0)$.

Conversely, if $\exists v \in C^\perp(XU) \cap C(XU_0)$.

then $v^T XU = 0$ $\therefore X^T v \perp C(U)$. $X^T v \in C(\Lambda)$

$\Rightarrow X^T v \in C(X^T) \cap C(\Lambda) = C(\Lambda_0) \perp C(U_0)$

$\therefore v \perp C(XU_0)$ i.e. $v = 0$. $\Rightarrow C^\perp(XU) \cap C(XU_0) = 0$

prop: If $\Lambda^T \beta$ is estimable. $\Lambda \neq 0$. Then $C(XU) \neq C(X)$.

Rmk: It means if estimable part isn't null, then it really reduces a model.

Pf: $0 = \Lambda^T U = P^T X U = P^T M X U$. $\therefore C(XU) \perp C(MP)$

By $C(XU)$. $C(MP) \subset C(X)$. $C(MP) \neq \{0\}$.

$\therefore C(XU) \subsetneqq C(X)$.

Cor. $C(\Lambda) \cap C(X^T) = 0 \Leftrightarrow C(XU) = C(X)$

Rmk: i) If $\Lambda^T \beta$ isn't estimable. then find Λ_0, $C(\Lambda_0) = C(X^T) \cap C(\Lambda)$, $\Lambda_0^T = P_0^T X$

And using $M - M_{MP_0}$

ii) $\Lambda^T \beta$ isn't estimable w.r.t each component

But joint constraint can generator an estimable constraint.

e.g. ANOVA $y_{ij} = \mu + \alpha_i + e_{ij}$

$\alpha_1 = 0$, $\alpha_2 = 0$ isn't estimable. But

$\alpha_1 - \alpha_2 = 0$ is estimable.

③ Generalized Hypothesis Test:

· Consider to test $H_0: \Lambda^T\beta = \lambda$. $\Lambda^T = P^T X$

$P^T \in M^{s\times n}$. $X \in M^{n\times p}$. $\lambda \in \mathbb{R}^s$. known.

suppose $b \in \mathbb{R}^s$ satisfies: $P^T X b = \lambda$

$\Rightarrow H_0: P^T X(\beta - b) = 0$. Set $\beta^* = \beta - b$

Then written $Y = X\beta + \varepsilon$. is:

$Y - Xb = X\beta^* + \varepsilon$ with $H_0: \Lambda^T\beta^* = 0$ --- (A)

i.e. $H_0: E(Y - Xb) \in C(X) \cap C^{\perp}(P) = C(X) \cap C^{\perp}(MP)$

Rmk: The reduced Model can be written as:

$$Y - Xb = X_0\gamma_0 + \varepsilon. \quad X_0 = M - M_{MP}.$$

prop. Test statistics is $F = \dfrac{\|M_{MP}(Y - Xb)\|^2 / r(M_{MP})}{\|(I-M)(Y-Xb)\|^2 / r(I-M)}$

$$= \frac{(\Lambda^T\hat\beta - \lambda)^T(\Lambda^T(X^TX)^-\Lambda)^-(\Lambda^T\hat\beta - \lambda) / r(\Lambda)}{\|(I-M)Y\|^2 / r(I-M)}$$

where $\Lambda^T = P^T X$. $\Lambda^T\hat\beta = P^T M Y$.

Rmk: It's indept with choice of b.

Thm. $F \sim F(r(M_{MP}), r(I-M), \gamma)$

$$= F(r(\Lambda), r(I-M), \gamma)$$

where $\gamma = \dfrac{\|(X\beta - Xb) M_{MP}\|^2}{\sigma^2} = \dfrac{(\Lambda^T\beta - \lambda)^T(\Lambda^T(X^TX)^-\Lambda)^-(\Lambda^T\beta - \lambda)}{\sigma^2}$

④ Breaking Sum of Squares into indept components:

· For $C(M_T) \subset C(M)$: $F = \dfrac{\|M_T Y\|^2 / r(M_T)}{\|(I-M)Y\|^2 / r(I-M)}$ defines

a test statistics for model (reduced):

$Y = (M - M_T)\gamma_0 + \varepsilon$. $\begin{cases} \text{estimation space}: C(M-M_T) \\ \text{error space}: C(I-(M-M_T)) \end{cases}$

Suppose $r(M_T) = r$. Decompose $C(M_T)$:

$C(M_T) = \sum_1^r C(M_i)$. $r(M_i) = 1$. $M_i M_j = 0$. $i \neq j$.

where M_i's are orthonormal proj's.

Thm. $R = (R_1 \cdots R_r)$. $\{R_i\}$ is orthonormal basis of $C(M_T)$. Then $M_T = RR^T = \sum_1^r R_i R_i^T$

Choose $M_i = R_i R_i^T$.

Rmk: $Y^T M_T Y = \sum_{i=1}^r Y^T M_i Y$. $F_i = \dfrac{\|M_i Y\|^2}{MSE}$

$\sim F(1, r(I-M), \dfrac{\|M_i X\beta\|^2}{\sigma^2})$

$F = \sum_1^r F_i$. Note $\|M X\beta\|^2 = 0 \iff$

$\|M_i X\beta\|^2 = 0$. $\forall 1 \le i \le r$. i.e. all H_0^i correspond M_i should be true.

⑤ Confidence Region:

Note that $\dfrac{\|M_{MP}(Y - X\beta)\|^2 / r(M_{MP})}{\|(I-M)(Y-X\beta)\|^2 / r(I-M)} =: F =$

$$(\Lambda^T\hat{\beta} - \Lambda^T\beta)^T(\Lambda^T(X^TX)^-\Lambda)^-(\Lambda^T\hat{\beta} - \Lambda^T\beta)/r(\Lambda)$$

$$\overline{\qquad\qquad MSE \qquad\qquad}$$

$\sim F(r(M_{MP}), r(I-M))$. Set $c_\alpha = F_{1-\alpha}(r(\Lambda), r(I-M))$

Obtain CI by $F \le c_\alpha$. (w.r.t. $\Lambda^T\beta$)

⑥ Likelihood Ratio Test:

For $H_0: E(Y) \in C(X_0)$ v.s. $H_1: E(Y) \in C(X) \cap C^\perp(X_0)$

Compute: $\lambda(Y) = \dfrac{\sup_{\theta_0} L(\theta|Y)}{\sup_{\theta} L(\theta|Y)}$. $R = \{Y \mid \lambda(Y) \le c\}$.

i) Numerator:

Under H_0: $Y = X_0\gamma_0 + \varepsilon$.

$$L = L(\gamma_0, \sigma^2|Y) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{(Y-X_0\gamma_0)^T(Y-X_0\gamma_0)}{2\sigma^2}}$$

Choose $\hat{\gamma}_0$ satisfies: $X_0\hat{\gamma}_0 = M_0Y$.

$$\hat{\sigma}^2 = \frac{Y^T(I-M_0)Y}{n} ;\quad M_0 = P_{C(X_0)|C^\perp(X_0)}$$

ii) Denominator:

$$L = L(\beta, \sigma^2|Y) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{(Y-X\beta)^T(Y-X\beta)}{2\sigma^2}}$$

Choose $\hat{\beta}$ satisfies $X\hat{\beta} = MY$.

$$\hat{\sigma}^2 = \frac{Y^T(I-M)Y}{n}$$

$$\Rightarrow \lambda(Y) = \left(\frac{Y^T(I-M)Y}{Y^T(I-M_0)Y}\right)^{\frac{n}{2}} \le c$$

equivalently. $\dfrac{Y^T(I-M)Y}{Y^T(I-M_0)Y} \le C_0 \iff \dfrac{Y^T(I-M_0)Y}{Y^T(I-M)Y} \ge C_1$

$\iff 1+\dfrac{Y^T(M-M_0)Y}{Y^T(I-M)Y} \ge C_1 \iff \dfrac{\|(M-M_0)Y\|^2/r(M-M_0)}{\|(I-M)Y\|^2/r(I-M)} \ge \tilde{C}$

Rmk: It's precise the test statistisc we have derived before.

(3) For generalized LSE:

Test $Y = X\beta + \varepsilon$. $\varepsilon \sim N_n(0, \sigma^2 V)$ v.s.

$Y = X_0\gamma + \varepsilon$, $C(X_0) \subset C(X)$. $V = QQ^T$

We gonna to reduce it to: $(*)$

$Q^{-1}Y = Q^{-1}X\beta + Q^{-1}\varepsilon$ v.s. $Q^{-1}Y = Q^{-1}X_0\gamma + Q^{-1}\varepsilon$.

Lemma. $C(X_0) \subset C(X)$. $|Q| \neq 0 \Rightarrow C(Q^{-1}X_0) \subset C(Q^{-1}X)$

Pf: $\forall\, v \in C(Q^{-1}X_0) \Rightarrow v = Q^{-1}X_0 b_1$

$X_0 = X b_2 \quad \therefore v = Q^{-1}X\tilde{b} \in C(Q^{-1}X)$

Thm. For $(*)$. The test statistics is

$$F = \frac{Y^T(A-A_0)^T V^{-1}(A-A_0)Y/(r(X)-r(X_0))}{Y^T(I-A)^T V^{-1}(I-A)Y/(n-r(X))}$$

$\sim F(r(X)-r(X_0),\ n-r(X),\ \gamma)$

$\gamma = \dfrac{\beta^T X^T(A-A_0)^T V^{-1}(A-A_0)X\beta}{\sigma^2}$. $\gamma = 0 \iff H_0$ holds.

Now, to test $H_0: \Delta^T\beta = 0$. $\Delta^T = P^T X$.

$$\Rightarrow F = \frac{\widehat{\beta^T\Delta} (\Delta^T (X^T V^{-1} X)^- \Delta)^- \widehat{\Delta^T\beta} / r(\Delta)}{MSE} \sim F(r(\Delta), n-r(X), \gamma)$$

$$\gamma = \frac{\beta^T\Delta (\Delta^T (X^T V^{-1} X)^- \Delta)^- \Delta^T\beta}{\sigma^2}, \quad \widehat{\Delta^T\beta} = P^T A Y.$$

Or we can rewrite it into:

$$F = \frac{Y^T A^T P (P^T X (X^T V^{-1} X)^- X^T P)^- P^T A Y}{MSE \cdot r(A^T P)} \sim F(r(A^T P), n-r(X), \gamma)$$

$$\gamma = \frac{\beta^T X^T P (P^T X (X^T V^{-1} X) X^T P)^- P^T X \beta}{\sigma^2}$$

Rmk: Obtain confidence interval as usual: consider

$$F = \frac{\|(M_{\tilde{m}P})(\tilde{Y} - \tilde{X}\beta)\|^2 / r(A^T)}{\|(I - M_{\tilde{X}})\tilde{Y}\|^2 / r(I - \tilde{M})}. \quad A^T = P^T X.$$

we get CI of $\Delta^T\beta$.