Recap: $\underline{y} = \underline{X}\,\underline{\beta} + \underline{e}$, $E[\underline{e}] = \underline{0}$, $V(\underline{e}) = \sigma^2 \underline{I}$

We started developing results when

$\underline{e} \sim N(\underline{0}, \sigma^2 \underline{I})$

$\frac{\underline{y}^T(\underline{I} - \underline{P})\underline{y}}{\sigma^2} \sim \chi^2(n - \text{rank}(\underline{X}))$

$\frac{\underline{y}^T \underline{P}\, \underline{y}}{\sigma^2} \sim \chi^2\left(\text{rank}(\underline{X}),\ \frac{\underline{\beta}^T \underline{X}^T \underline{X}\, \underline{\beta}}{2\sigma^2}\right)$

$\underline{y}^T \underline{P}\,\underline{y}$ and $\underline{y}^T(\underline{I} - \underline{P})\underline{y}$ are independent.

What is the dist. of

$$\frac{\dfrac{\underline{y}^T P \underline{y}}{\text{rank}(\underline{X})}}{\dfrac{\underline{y}^T(\underline{I} - \underline{P})\underline{y}}{n - \text{rank}(\underline{X})}} \quad ?$$

$\underline{y}^T P \underline{y} = \|\hat{\underline{y}}\|^2$ as $\hat{\underline{y}} = P\underline{y}$

$\underline{y}^T(\underline{I} - \underline{P})\underline{y} = \|\hat{\underline{e}}\|^2$ as $\hat{\underline{e}} = (\underline{I} - \underline{P})\underline{y}$

$$= \frac{\dfrac{\|\hat{\underline{y}}\|^2}{\text{rank}(\underline{X})}}{\dfrac{\|\hat{\underline{e}}\|^2}{n - \text{rank}(\underline{X})}}$$

Def: Let u_1 and u_2 be independent random variables, with $u_1 \sim \chi^2_{p_1}$ and $u_2 \sim \chi^2_{p_2}$

then $f = \frac{u_1/p_1}{u_2/p_2} \sim F_{p_1, p_2}$ (F dist. p_1, p_2 degrees of freedom).

Def:- Let u_1 and u_2 be independent random variables, with $u_1 \sim \chi^2_{p_1, \phi}$ and $u_2 \sim \chi^2_{p_2}$

then $f = \dfrac{u_1/p_1}{u_2/p_2}$ has the noncentral F distribution with degrees of freedom p_1 and p_2 and noncentrality parameter ϕ, denoted by $f \sim F_{p_1, p_2, \phi}$.

$$\frac{\underline{y}^T \underline{P}\, \underline{y} \,/\, \text{rank}(\underline{X})}{\underline{y}^T(\underline{I} - \underline{P})\,\underline{y} \,/\, (n - \text{rank}(\underline{X}))} \sim F_{\text{rank}(\underline{X}),\, n-\text{rank}(\underline{X}),\, \frac{\underline{\beta}^T \underline{X}^T \underline{X}\, \underline{\beta}}{2\sigma^2}}$$

Def:- Let $u \sim N(\mu, 1)$ and $v \sim \chi^2_K$. If u and v are independent, then $t = \dfrac{u}{\sqrt{\frac{v}{K}}}$ has the noncentered students t-distribution with K degrees of freedom and a noncentrality parameter μ, denoted by $t \sim t_{K,\mu}$. When $\mu = 0$, we get back to the standard students t-dist.

Let $\underline{\lambda}^T \underline{\beta}$ be any estimable function. $\underline{\lambda}^T \hat{\underline{\beta}}$ is the unbiased estimator of $\underline{\lambda}^T \underline{\beta}$.

$\cancel{x^T\beta}$ what is the dist. of $\lambda^T\hat{\beta}$.

$$\underline{\lambda}^T\hat{\underline{\beta}} \sim N\left(\underline{\lambda}^T\underline{\beta},\ \sigma^2\underline{\lambda}^T(\underline{X}^T\underline{X})^-\underline{\lambda}\right)$$

$$\frac{\underline{\lambda}^T\hat{\underline{\beta}} - \underline{\lambda}^T\underline{\beta}}{\sqrt{\sigma^2\lambda^T(\underline{X}^T\underline{X})^-\underline{\lambda}}} \sim N(0,1)$$

$$\frac{\underline{y}^T(\underline{I}-\underline{P})\underline{y}}{\sigma^2} \sim \chi^2_{n-\text{rank}(\underline{X})}$$

$$\Rightarrow \frac{\dfrac{\underline{\lambda}^T\hat{\underline{\beta}} - \underline{\lambda}^T\underline{\beta}}{\sqrt{\sigma^2\lambda^T(\underline{X}^T\underline{X})^-\underline{\lambda}}}}{\sqrt{\dfrac{\underline{y}^T(\underline{I}-\underline{P})\underline{y}}{\sigma^2(n-\text{rank}(\underline{X}))}}} \sim t_{n-\text{rank}(\underline{X})}$$

Now note that,

$$\frac{\dfrac{\underline{\lambda}^T\hat{\underline{\beta}}}{\sqrt{\sigma^2\underline{\lambda}^T(\underline{X}^T\underline{X})^-\underline{\lambda}}}}{\sqrt{\underbrace{\dfrac{\underline{y}^T(\underline{I}-\underline{P})\underline{y}}{\sigma^2(n-\text{rank}(\underline{X}))}}_{\frac{\chi^2_{n-\text{rank}(\underline{X})}}{n-\text{rank}(\underline{X})}}}} \sim t_{n-\text{rank}(\underline{X}),\ \lambda^T\underline{\beta}}$$

(numerator) $\sim$ Normal with a ~~different~~ mean different from 0

One way Anova Model:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \qquad \begin{array}{l} i = 1, \ldots, a, \\ j = 1, \ldots, n_i \end{array}$$

$$\underline{y} = \underbrace{\begin{pmatrix} y_{11} \\ \vdots \\ \vdots \\ y_{an_a} \end{pmatrix}}_{\underline{y}} = \underbrace{\begin{bmatrix} \underbrace{\begin{matrix} \underline{1}_{n_1} \\ \underline{1}_{n_2} \\ \vdots \\ \vdots \\ \underline{1}_{n_a} \end{matrix}}_{\underline{W}} & \begin{matrix} \underline{1}_{n_1} \\ \underline{0}_{n_2} \\ \vdots \\ \vdots \\ \underline{0}_{n_a} \end{matrix} & \cdots\cdots & \begin{matrix} \underline{0}_{n_1} \\ \vdots \\ \vdots \\ \underline{0}_{n_{a-1}} \\ \underline{1}_{n_a} \end{matrix} \end{bmatrix}}_{\underline{X}} \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \vdots \\ \alpha_a \end{pmatrix} + \underline{e}$$

$$\begin{aligned} P_W &= \underline{W} \left(\underline{W}^T \underline{W} \right)^{-1} \underline{W}^T \\ &= \underline{1}_{n_1+\cdots+n_a} \left(\underline{1}^T_{n_1+\cdots+n_a} \, \underline{1}_{n_1+\cdots+n_a} \right)^{-1} \underline{1}^T_{n_1+\cdots+n_a} \\ &= \frac{1}{n_1+\cdots+n_a} \underline{1}_{n_1+\cdots+n_a} \underline{1}^T_{n_1+\cdots+n_a} \\ &= \frac{1}{n_1+\cdots+n_a} J_{n_1+\cdots+n_a}. \end{aligned}$$

$$\underline{P_W} \, \underline{y} = \frac{1}{(n_1+\cdots+n_a)} J_{n_1+\cdots+n_a} \, \underline{y} = \begin{pmatrix} \bar{y}_{..} \\ \vdots \\ \vdots \\ \bar{y}_{..} \end{pmatrix}$$

$\bar{y}_{..}$ = average of all y_{ij}'s.

$C(\underline{W}) \subset C(\underline{X})$.

$$\underset{\sim}{P}_X = \underset{\sim}{X} \left(\underset{\sim}{X}^T \underset{\sim}{X} \right)^{-} \underset{\sim}{X}$$

$$= \begin{bmatrix} \frac{1}{n_1} \underset{\sim}{J}_{n_1} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \frac{1}{n_2} \underset{\sim}{J}_{n_2} & & \vdots \\ \vdots & & \ddots & \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \frac{1}{n_a} \underset{\sim}{J}_{n_a} \end{bmatrix}$$

$$\underset{\sim}{P}_X \underset{\sim}{y} = \begin{bmatrix} \frac{1}{n_1} \underset{\sim}{J}_{n_1} & & \\ & \ddots & \\ & & \frac{1}{n_a} \underset{\sim}{J}_{n_a} \end{bmatrix} \begin{pmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{a1} \\ \vdots \\ y_{an_a} \end{pmatrix}$$

$$= \begin{pmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \vdots \\ \bar{y}_{2\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{pmatrix} \begin{matrix} \} \; n_1 \text{ times} \\ \\ \} \; n_2 \text{ times} \\ \\ \} \; n_a \text{ times} \end{matrix}$$

$\bar{y}_{1\cdot}$ = average of group 1

$\vdots$

$\bar{y}_{a\cdot}$ = average of group a.

$$\left(\underset{\sim}{P}_X - \underset{\sim}{P}_W \right) \underset{\sim}{y} = \begin{pmatrix} \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{a\cdot} - \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{a\cdot} - \bar{y}_{\cdot\cdot} \end{pmatrix} \begin{matrix} \} \; n_1 \text{ times} \\ \vdots \\ \} \; n_a \text{ times} \end{matrix}$$

$$(\underline{I} - \underline{P}_X)\underline{y} = \begin{pmatrix} y_{11} - \bar{y}_{1.} \\ \vdots \\ y_{1n_1} - \bar{y}_{1.} \\ \vdots \\ y_{a1} - \bar{y}_{a.} \\ \vdots \\ y_{an_a} - \bar{y}_{a.} \end{pmatrix}$$

$$\underline{y}^T\underline{y} = \underline{y}^T\left[P_W + P_X - P_W + \underline{I} - \underline{P}_X\right]\underline{y}$$

$$= \underline{y}^T P_W \underline{y} + \underline{y}^T(\underline{P}_X - \underline{P}_W)\underline{y} + \underline{y}^T(\underline{I} - \underline{P}_X)\underline{y}$$

Is P_W idempotent? Yes.

Is $(\underline{I} - P_X)$ idempotent? Yes

Is $(P_X - P_W)$ idempotent? Yes. $(\underline{P}_X - \underline{P}_W)(P_X - P_W)$

$$= P_X^2 - P_X P_W - P_W P_X + P_W^2$$

$$= P_X - P_W - \cancel{P_W} + \cancel{P_W}$$

Thm: (Cochran's theorem).

Let $\underline{y} \sim N(\underline{\mu}, \sigma^2 \underline{I})$ and let $\underline{A}_i$, $i = 1, \ldots, K$ be symmetric idempotent matrices with rank s_i. If $\sum_{i=1}^{K} A_i = \underline{I}$, then, $\dfrac{\underline{y}^T \underline{A}_i \underline{y}}{\sigma^2}$ are independently dist. as $\chi^2_{s_i, \phi_i}$ where $\phi_i = \frac{1}{2\sigma^2} \underline{\mu}' \underline{A}_i \mu$. and $\sum_{i=1}^{K} s_i = n$.

In our case, $\underline{P}_W$, $\underline{P}_X - \underline{P}_W$ and $\underline{I} - \underline{P}_X$ all three are idempotent and symmetric

$\underline{y} \sim N(\underline{X}\underline{\beta}, \sigma^2 \underline{I})$ and $\underline{P}_W + \underline{P}_X - \underline{P}_W + \underline{I} - \underline{P}_X = \underline{I}$.

thus, $$\frac{\underline{y}^T \underline{P}_W \underline{y}}{\sigma^2} \sim \chi^2_{1,\ \frac{(\underline{X}\underline{\beta})' \underline{P}_W (\underline{X}\underline{\beta})}{2\sigma^2}}$$

$$\frac{\underline{y}^T (\underline{P}_X - \underline{P}_W) \underline{y}}{\sigma^2} \sim \chi^2_{\text{rank}(\underline{X})-1,\ \frac{(\underline{X}\underline{\beta})' (\underline{P}_X - \underline{P}_W)(\underline{X}\underline{\beta})}{2\sigma^2}}$$

$$\frac{\underline{y}^T (\underline{I} - \underline{P}_X) \underline{y}}{\sigma^2} \sim \chi^2_{n-\text{rank}(\underline{X}),\ \underbrace{\frac{(\underline{X}\underline{\beta})' (\underline{I} - \underline{P}_X) \underline{X}\underline{\beta}}{2\sigma^2}}_{0}}$$

Anova Table

$n = n_1 + \cdots + n_a$

Source	df	SS
Mean	1	$\underline{y}^T \underline{P}_W \underline{y} = n \bar{y}_{..}^2$
Group	$a-1$	$\underline{y}^T (\underline{P}_X - \underline{P}_W) \underline{y} = \sum_{i=1}^{a} n_i (\bar{y}_{i.} - \bar{y}_{..})^2$
Error	$n-a$	$\underline{y}^T (\underline{I} - \underline{P}_X) \underline{y} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$

$$\underline{y}^T \underline{P}_W \underline{y} = (\underline{P}_W \underline{y})^T (\underline{P}_W \underline{y}) = (\bar{y}_{..} \cdots\cdots \bar{y}_{..}) \begin{pmatrix} \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \end{pmatrix}$$

$$= n \bar{y}_{..}^2$$

$$\underline{y}^T(\underline{P}_X - \underline{P}_W)\,\underline{y} = \left[(\underline{P}_X - \underline{P}_W)\,\underline{y}\right]^T \left[(\underline{P}_X - \underline{P}_W)\,\underline{y}\right]$$

$$= \Big(\underbrace{\bar{y}_{1.} - \bar{y}_{..}, \ldots, \bar{y}_{1.} - \bar{y}_{..}}_{n_1 \text{ times}}, \ldots\ldots, \underbrace{\bar{y}_{a.} - \bar{y}_{..}, \ldots, \bar{y}_{a.} - \bar{y}_{..}}_{n_a \text{ times}}\Big) \begin{pmatrix} \bar{y}_{1.} - \bar{y}_{..} \\ \vdots \\ \bar{y}_{1.} - \bar{y}_{..} \\ \vdots \\ \bar{y}_{a.} - \bar{y}_{..} \\ \vdots \\ \bar{y}_{a.} - \bar{y}_{..} \end{pmatrix}$$

$$= \sum_{i=1}^{a} n_i \left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$\underline{y}^T(\underline{I} - \underline{P}_X)\,\underline{y}$$

$$= \left[(\underline{I} - \underline{P}_X)\,\underline{y}\right]^T \left[(\underline{I} - \underline{P}_X)\,\underline{y}\right]$$

$$= \left(y_{11} - \bar{y}_{1.}, \ldots, y_{1n_1} - \bar{y}_{1.}, \ldots, y_{a1} - \bar{y}_{a.}, \ldots, y_{an_a} - \bar{y}_{a.}\right) \begin{pmatrix} y_{11} - \bar{y}_{1.} \\ \vdots \\ y_{1n_1} - \bar{y}_{1.} \\ \vdots \\ y_{a1} - \bar{y}_{a.} \\ \vdots \\ y_{an_a} - \bar{y}_{a.} \end{pmatrix}$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_{i.}\right)^2$$

Compute noncentrality parameters

$$\frac{(\underline{X}\underline{\beta})'\,\underline{P}_W\,(\underline{X}\underline{\beta})}{2\sigma^2} = \frac{(\underline{P}_W\,\underline{X}\underline{\beta})'(\underline{P}_W\,\underline{X}\underline{\beta})}{2\sigma^2}$$

$$\underline{P}_W \underline{X}\,\underline{\beta} = \frac{1}{(n_1+\cdots+n_a)} J_{n_1+\cdots+n_a} \begin{bmatrix} \mu+\alpha_1 \\ \vdots \\ \mu+\alpha_1 \\ \vdots \\ \mu+\alpha_a \\ \vdots \\ \mu+\alpha_a \end{bmatrix} \begin{matrix} \} \ n_1 \text{ times} \\ \\ \\ \\ \} \ n_a \text{ times}. \end{matrix}$$

$$= \begin{bmatrix} \mu + \dfrac{\sum_{i=1}^{a} n_i\alpha_i}{\sum_{i=1}^{a} n_i} \\ \vdots \\ \mu + \dfrac{\sum_{i=1}^{a} n_i\alpha_i}{\sum_{i=1}^{a} n_i} \end{bmatrix}$$

$$(\underline{P}_W \underline{X}\,\underline{\beta})'(\underline{P}_W \underline{X}\,\underline{\beta}) = (n_1+\cdots+n_a)\left(\mu + \frac{n_1\alpha_1+\cdots+n_a\alpha_a}{n_1+\cdots+n_a}\right)^2$$

$$(X\underline{\beta})'(\underline{P}_X - \underline{P}_W)\underline{X}\,\underline{\beta} = \left[(P_X - P_W)\underline{X}\,\underline{\beta}\right]'\left[(P_X - P_W)\underline{X}\,\underline{\beta}\right]$$

$$(\underline{P}_X - \underline{P}_W)\underline{X}\,\underline{\beta} = (\underline{P}_X - \underline{P}_W) \begin{bmatrix} \mu+\alpha_1 \\ \vdots \\ \mu+\alpha_1 \\ \vdots \\ \mu+\alpha_a \\ \vdots \\ \mu+\alpha_a \end{bmatrix} \begin{matrix} \} \ n_1 \text{ times} \\ \\ \\ \\ \} \ n_a \text{ times}. \end{matrix}$$

$$\left[(P_X - P_W)\underline{X}\,\underline{\beta}\right]'\left[(P_X - P_W)\underline{X}\,\underline{\beta}\right] = \sum_{i=1}^{a} n_i(\alpha_i - \bar{\alpha})^2, \quad \bar{\alpha} = \frac{\sum_{i=1}^{a} n_i\alpha_i}{\sum_{i=1}^{a} n_i}$$

$$\frac{\underline{y}^T P_W \underline{y}}{\sigma^2} = \frac{n\,\bar{y}_{..}^2}{\sigma^2} \sim \chi^2_{1,\ \left(\mu + \frac{\sum_{i=1}^{a} n_i \alpha_i}{\sum_{i=1}^{a} n_i}\right)^2 \frac{\sum_{i=1}^{a} n_i}{2\sigma^2}}$$

$$\frac{\underline{y}^T (P_X - P_W)\underline{y}}{\sigma^2} = \frac{\sum_{i=1}^{a} n_i (\bar{y}_{i.} - \bar{y}_{..})^2}{\sigma^2}$$

$$\sim \chi^2_{a-1,\ \frac{\sum_{i=1}^{a} n_i (\alpha_i - \bar{\alpha})^2}{2\sigma^2}}$$

$$\frac{\underline{y}^T (\underline{I} - \underline{P}_X)\underline{y}}{\sigma^2} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{\sigma^2} \sim \chi^2_{n-a}$$

Source	df	SS	Noncentrality	$= \frac{SS}{df}$ MS
Mean	1	$n\bar{y}_{..}^2$	$\left(\mu + \frac{\sum_{i=1}^{a} n_i\alpha_i}{\sum_{i=1}^{a} n_i}\right)^2 \frac{\sum_{i=1}^{a} n_i}{2\sigma^2}$	$n\bar{y}_{..}^2$
Group	$a-1$	$\sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2$	$\frac{\sum_{i=1}^{a} n_i(\alpha_i - \bar{\alpha})^2}{2\sigma^2}$	$\frac{\sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2}{a-1}$
Error	$n-a$	$\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2$	0	$\frac{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}{n-a}$

Recall: We studied one-way-Anova, Cochran theorem. Determine distributions of $\underline{y}^T \underline{A}_i \underline{y}$, where $\underline{A}_i$ is an idempotent matrix.

Now, when $\underline{y} \sim N(\underline{X}\,\underline{\beta}, \sigma^2 \underline{I})$

① $\underline{\Lambda}^T \hat{\underline{\beta}} \sim N(\underline{\Lambda}^T \underline{\beta}, \sigma^2 \underline{\Lambda}^T (\underline{X}^T \underline{X})^{-} \underline{\Lambda})$ where $\underline{\Lambda}$ is a $p \times s$ matrix, $\underline{\beta}$ is a $p \times 1$ dimensional vector.

② $\frac{SSE}{\sigma^2} = \frac{\underline{y}^T (\underline{I} - \underline{P}) \underline{y}}{\sigma^2} \sim \chi^2_{n-\text{rank}(\underline{X})}$

③ $\underline{\Lambda}^T \hat{\underline{\beta}}$ and $\frac{SSE}{\sigma^2}$ are independent.

Testing:

① Testing linear parametric functions, i.e.
$H_0: \underline{\Lambda}^T \underline{\beta} = \underline{m}$ vs. $H_1: \underline{\Lambda}^T \underline{\beta} \neq \underline{m}$.

② Testing a reduced model.

an example of ② is the following:
Consider the one way anova model

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \dots, a, \; j = 1, \dots, n_i.$$

the above model let be called as the full model.

$H_0: \alpha_1 = \cdots = \alpha_a$

$$y_{ij} = \tilde{\mu} + \tilde{e}_{ij}, \quad i = 1, \dots, a, \; j = 1, \dots, n_i$$

this is an example of a reduced model.

①

Generally $H_0: \underline{\Lambda}^T \underline{\beta} = \underline{m}$ vs. $H_1: \underline{\Lambda}^T \underline{\beta} \neq \underline{m}$.

$\underline{\Lambda}$ is a $p \times s$ matrix with a full column rank.

$$\underline{\Lambda}_{p \times s} \qquad \underline{\Lambda} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1s} \\ \vdots & & \vdots \\ \lambda_{p1} & \cdots & \lambda_{ps} \end{pmatrix}$$

$$\underline{\Lambda}^T \underline{\beta} = \underline{m} \Rightarrow \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{p1} \\ \vdots & & \vdots \\ \lambda_{1s} & \cdots & \lambda_{ps} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_s \end{pmatrix}$$

$$\Rightarrow \begin{array}{c} \lambda_{11}\beta_1 + \cdots + \lambda_{p1}\beta_p = m_1 \\ \vdots \\ \lambda_{1s}\beta_1 + \cdots + \lambda_{ps}\beta_p = m_s \end{array}$$

Also what we require is that each component of

$$\underline{\Lambda}^T \underline{\beta} = \begin{pmatrix} \lambda_{11}\beta_1 + \cdots + \lambda_{p1}\beta_p \\ \vdots \\ \lambda_{1s}\beta_1 + \cdots + \lambda_{ps}\beta_p \end{pmatrix} = \begin{pmatrix} \underline{\lambda}_1^T \underline{\beta} \\ \vdots \\ \underline{\lambda}_s^T \underline{\beta} \end{pmatrix}$$

each component $\underline{\lambda}_j^T \underline{\beta}$ is an estimable function of $\underline{\beta}$.

Def:- The general linear hypothesis $H_0: \underline{\Lambda}^T \underline{\beta} = \underline{m}$ is testable if and only if $\underline{\Lambda}$ has full column rank and each component of $\underline{\Lambda}^T \underline{\beta}$ is estimable.

Recall that,

$$\underline{\Lambda}^T \hat{\underline{\beta}} \sim N\left(\underline{\Lambda}^T \underline{\beta}, \ \sigma^2 \underbrace{\underline{\Lambda}^T (\underline{X}^T \underline{X})^- \underline{\Lambda}}_{\underline{H}}\right)$$

$$\Rightarrow \underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m} \sim N_s\left(\underline{\Lambda}^T \underline{\beta} - \underline{m}, \ \sigma^2 \underline{H}\right)$$

$$\Rightarrow \left(\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m}\right)' \left(\sigma^2 \underline{H}\right)^{-1} \left(\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m}\right)$$

$$\sim \chi^2_{s, \phi} \quad \text{- - - - - - - (1)}$$

$$\phi = \frac{1}{2} \left(\underline{\Lambda}^T \underline{\beta} - \underline{m}\right)' \left(\sigma^2 \underline{H}\right)^{-1} \left(\underline{\Lambda}^T \underline{\beta} - \underline{m}\right)$$

under $H_0: \underline{\Lambda}^T \underline{\beta} = \underline{m} \Rightarrow \phi = 0$

Now note that,

$SSE = \underline{y}^T (\underline{I} - \underline{P}) \underline{y}$ is independent of

$\underline{\Lambda}^T \hat{\underline{\beta}}$ and $\frac{SSE}{\sigma^2} \sim \chi^2_{n - \text{rank}(\underline{X})}$ - - - (2)

~~Dividing (1) by (2) $\frac{SSE}{(n-}$~~

$$\frac{\left(\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m}\right)' \left(\sigma^2 \underline{H}\right)^{-1} \left(\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m}\right) / s}{\frac{1}{\sigma^2} \frac{SSE}{n - \text{rank}(\underline{X})}}$$

$$= \frac{\chi^2_{s, \phi} / s}{\frac{\chi^2_{n - \text{rank}(\underline{X})}}{n - \text{rank}(\underline{X})}} \sim F_{s, \, n - \text{rank}(\underline{X}), \, \phi}$$

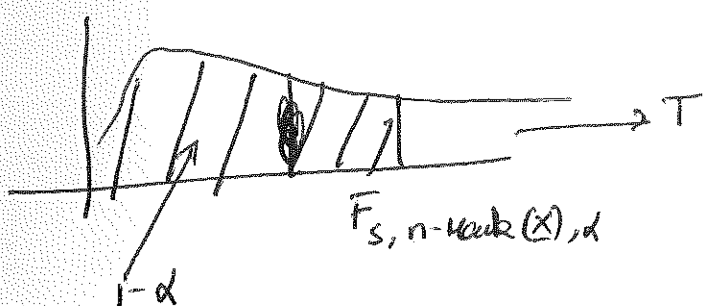
under H_0, $\phi = 0$

$\Rightarrow$ under H_0,

$$T = \frac{\dfrac{(\underset{\sim}{\Lambda}^T \hat{\underset{\sim}{\beta}} - \underset{\sim}{m})' H^{-1} (\underset{\sim}{\Lambda}^T \hat{\underset{\sim}{\beta}} - \underset{\sim}{m})}{s}}{\dfrac{SSE}{n - \text{rank}(\underset{\sim}{X})}} \sim F_{s,\, n-\text{rank}(\underset{\sim}{X})}$$

Given data, you will compute T and you reject the null hypothesis H_0 if

$$T > F_{s,\, n-\text{rank}(\underset{\sim}{X}),\, \alpha}$$



when $s = 1$, i.e. there is only one linear hypothesis that you are testing $H_0: \underset{\sim}{\lambda}^T \underset{\sim}{\beta} = m$

$$\underset{\sim}{\lambda}^T \hat{\underset{\sim}{\beta}} - m \sim N\left(\underset{\sim}{\lambda}^T \underset{\sim}{\beta} - m,\ \sigma^2 \underset{\sim}{\lambda}^T (\underset{\sim}{X}^T \underset{\sim}{X})^- \underset{\sim}{\lambda}\right)$$

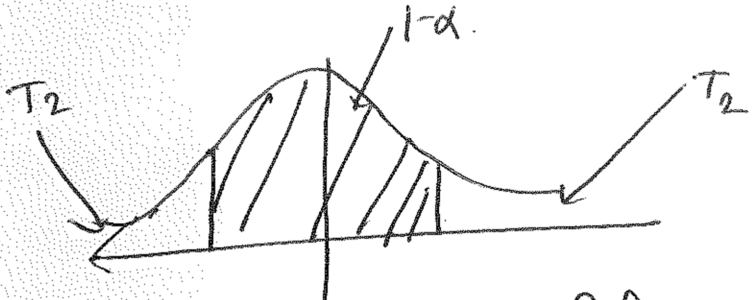
$$T_2 = \frac{\dfrac{\lambda^T \hat{\underset{\sim}{\beta}} - m}{\sqrt{\sigma^2 \lambda^T (X^T \underset{\sim}{X})^- \underset{\sim}{\lambda}}}}{\sqrt{\dfrac{SSE}{\sigma^2 (n - \text{rank}(\underset{\sim}{X}))}}} \overset{\text{under } H_0}{\sim} t_{n-\text{rank}(\underset{\sim}{X})}$$

$$T_2 = \frac{\dfrac{\lambda^T \hat{\underline{\beta}} - m}{\sqrt{\lambda^T (\underline{X}^T \underline{X})^- \underline{\lambda}}}}{\sqrt{\dfrac{SSE}{n - \text{rank}(\underline{X})}}} \overset{\text{under } H_0}{\sim} t_{n-\text{rank}(\underline{X})}$$

$$T_2^2 = \frac{(\underline{\lambda}^T \hat{\underline{\beta}} - m)^2}{\left[\lambda^T (\underline{X}^T \underline{X})^- \underline{\lambda}\right]\left[\dfrac{SSE}{n-\text{rank}(\underline{X})}\right]} \overset{\text{under } H_0}{\sim} F_{1,\, n-\text{rank}(\underline{X})}$$

the null hypothesis is rejected.

$$|T_2| > t_{n-\text{rank}(\underline{X}),\, \frac{\alpha}{2}}$$



<u>Testing reduced models</u>

start with the linear regression model

$\underline{y} = \underline{X}\,\underline{\beta} + \underline{e}, \quad \underline{e} \sim N(\underline{0}, \sigma^2 \underline{I}_n)$

Our wish is to use a more simple model

$\underline{y} = \underline{X}_0\,\underline{\gamma} + \underline{e}, \quad \underline{e} \sim N(\underline{0}, \sigma^2 \underline{I}_n)$

by simpler model we mean $C(\underline{X}_0) \subset C(\underline{X})$.

Ex: One way Anova.

$y_{ij} = \mu + \alpha_i + e_{ij}$, $i = 1, 2, 3,$ and $j = 1, 2$.

$\underline{y} = \underline{X}\,\underline{\beta} + \underline{e}$, $\underline{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$, $\underline{\beta} = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$

Suppose, I want to test if all the factors lead to similar response.

$\underline{y} = \underline{X}_0\,\gamma + \underline{e}$, $\underline{X}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$C(\underline{X}_0) \subset C(\underline{X})$

thus the second model is thought of as a reduced model to the earlier one.

Ex: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + e_i$, $i = 1, \ldots, n$.

$\underline{y} = \underline{X}\,\underline{\beta} + \underline{e}$, $\underline{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & x_{3n} \end{bmatrix}$, $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$

Suppose I would like to check if the 2nd and third predictors are explaining the response similarly $\Leftrightarrow \beta_2 = \beta_3$.

If I use $\beta_2 = \beta_3$

Then the reduced model will be

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 (x_{2i} + x_{3i}) + e_i$$

$$\underline{y} = \underline{X}_0 \underline{\gamma} + \underline{e}, \quad \underline{X}_0 = \begin{bmatrix} 1 & x_{11} & x_{21}+x_{31} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n}+x_{3n} \end{bmatrix}, \quad \underline{\gamma} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

$C(\underline{X}_0) \subset C(\underline{X})$

this is a legitimate reduced model.

We start with $\underline{y} = \underline{X}\,\underline{\beta} + \underline{e}$ i.e. $E[\underline{y}] = \underline{X}\,\underline{\beta}$

$\Leftrightarrow E[\underline{y}] \in C(\underline{X})$

$H_0 : E[\underline{y}] \in C(\underline{X}_0)$

$H_1 : E[\underline{y}] \in C(\underline{X}), \quad E[\underline{y}] \notin C(\underline{X}_0)$.

$\underline{Qn}$: How to build a test statistic for testing this?

Let $\underline{P}$ and $\underline{P}_0$ be projection matrices onto $C(\underline{X})$ and $C(\underline{X}_0)$ respectively.

$\hat{\underline{y}}$ = projected $\underline{y}$ onto the $C(\underline{X}) = \underline{P}\,\underline{y}$

$\hat{\hat{\underline{y}}}$ = projected $\underline{y}$ onto the $C(\underline{X}_0) = \underline{P}_0\,\underline{y}$

If the reduced model is close to the ~~[illegible]~~ full model then, $\|\underline{P}\underline{y} - \underline{P}_0 \underline{y}\|^2$ should be small

$$\|\underline{P}\underline{y} - \underline{P}_0 \underline{y}\|^2 = (\underline{P}\underline{y} - \underline{P}_0 \underline{y})^T (\underline{P}\underline{y} - \underline{P}_0 \underline{y})$$

$$= \underline{y}^T (\underline{P} - \underline{P}_0)^T (\underline{P} - \underline{P}_0)\, \underline{y}$$

$$= \underline{y}^T (\underline{P} - \underline{P}_0)\, \underline{y} \quad \left[\underline{P} - \underline{P}_0 \text{ is symmetric and idempotent} \right]$$

$$\frac{\underline{y}^T(\underline{P} - \underline{P}_0)\underline{y}}{\sigma^2} \sim \chi^2_{\text{rank}(\underline{P}-\underline{P}_0),\ \frac{(\underline{X}\underline{\beta})'(\underline{P}-\underline{P}_0)(\underline{X}\underline{\beta})}{2\sigma^2}}$$

Also, $\dfrac{\underline{y}^T(\underline{I} - \underline{P})\underline{y}}{\sigma^2} \sim \chi^2_{n-\text{rank}(\underline{X})}$. $\quad [(\underline{P}-\underline{P}_0)(\underline{I}-\underline{P}) = 0]$

thus, $\underline{y}^T(\underline{P} - \underline{P}_0)\underline{y}$ and $\underline{y}^T(\underline{I} - \underline{P})\underline{y}$ are independent

$$\Rightarrow \quad M = \frac{\dfrac{\underline{y}^T(\underline{P} - \underline{P}_0)\underline{y}}{\sigma^2\, \text{rank}(\underline{P}-\underline{P}_0)}}{\dfrac{\underline{y}^T(\underline{I} - \underline{P})\underline{y}}{\sigma^2(n - \text{rank}(\underline{X}))}} \sim F_{\text{rank}(\underline{P}-\underline{P}_0),\ n-\text{rank}(\underline{X}),\ \phi}$$

$$\phi = \frac{(\underline{X}\underline{\beta})'(\underline{P}-\underline{P}_0)(\underline{X}\underline{\beta})}{2\sigma^2}$$

when H_0 is true, $E[\underline{y}] \in C(\underline{X}_0)$

$\Rightarrow \phi = 0$

$$\text{thus,} \quad M = \frac{\dfrac{\underline{y}^T(\underline{P} - \underline{P}_0)\underline{y}}{\text{rank}(\underline{P}-\underline{P}_0)}}{\dfrac{\underline{y}^T(\underline{I} - \underline{P})\underline{y}}{n - \text{rank}(\underline{X})}} \overset{\text{under } H_0}{\sim} F_{\text{rank}(\underline{P}-\underline{P}_0),\ n-\text{rank}(\underline{X})}.$$

the reduced model is not supported by the data

if $M > F_{\text{rank}(\underline{P} - \underline{P}_0),\ n - \text{rank}(\underline{X}),\ \alpha}$.

The above procedure gives us a $100(1-\alpha)\%$ test for testing the reduced model.

The above ~~test statistic can be derived also from~~

<u>Testing</u>: ① Testing of linear parametric functions $H_0: \underline{\Lambda}^T \underline{\beta} = \underline{m}$

To test this, first check testability of hypotheses.

then the test statistic was also given to test the hypotheses.

if $s = 1$, $H_0: \underline{\lambda}^T \underline{\beta} = m$

then we can use t-test statistic.

② Checking the legitimacy of a reduced model. That is checked by the test statistic M.