

Recap: $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $i=1, \dots, a$; $j=1, \dots, n_i$

$$\underline{y} = \begin{pmatrix} y_{11} \\ \vdots \\ y_{an_a} \end{pmatrix} = \otimes \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{1}_{n_2} & \mathbf{0}_{n_3} & \cdots & \mathbf{0}_{n_a} \\ \vdots & \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \cdots & \mathbf{0}_{n_a} \\ \mathbf{1}_{n_a} & \mathbf{0}_{n_a} & \cdots & \cdots & \mathbf{1}_{n_a} \end{bmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \end{pmatrix} + \underline{\epsilon}$$

$$P_W = \mathbf{1}_{n_1+\cdots+n_a} \left(\mathbf{1}_{n_1+n_a}^T \mathbf{1}_{n_1+\cdots+n_a} \right)^{-1} \mathbf{1}_{n_1+\cdots+n_a}$$

$$= \frac{1}{(n_1+\cdots+n_a)} \mathbf{J}_{n_1+\cdots+n_a}$$

$$\underbrace{P_W \underline{y}}_{(n_1+\cdots+n_a) \times 1} = \begin{pmatrix} \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \end{pmatrix} \quad \bar{y}_{..} = \text{average of all } y_{ij}'s.$$

$C(W) \subset C(C(\underline{x}))$,

$$P_X = \underline{x} (\underline{x}^T \underline{x})^{-1} \underline{x}^T = \begin{bmatrix} \frac{1}{n_1} \mathbf{J}_{n_1} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{n_2} \mathbf{J}_{n_2} & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \cdots & \frac{1}{n_a} \mathbf{J}_{n_a} \end{bmatrix}$$

$$\underbrace{P_X \underline{y}}_{(n_1+\cdots+n_a) \times 1} = \begin{pmatrix} \bar{y}_{..} \\ \bar{y}_{..} \\ \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \\ \bar{y}_{..} \end{pmatrix} \quad \left\{ \begin{array}{l} \bar{y}_{..} \\ \bar{y}_{..} \\ \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \\ \bar{y}_{..} \end{array} \right\}_{n_1} \quad \left\{ \begin{array}{l} \bar{y}_{..} \\ \bar{y}_{..} \\ \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \\ \bar{y}_{..} \end{array} \right\}_{n_2} \quad \left\{ \begin{array}{l} \bar{y}_{..} \\ \bar{y}_{..} \\ \bar{y}_{..} \\ \vdots \\ \bar{y}_{..} \\ \bar{y}_{..} \end{array} \right\}_{n_a}$$

$P_X - P_W$ is the projection matrix onto the orthogonal space of the $C(W)$ in $C(\underline{x})$.

$$(P_X - P_W) \underline{y} = \begin{pmatrix} \bar{y}_{11} - \bar{y}_{..} \\ \bar{y}_{12} - \bar{y}_{..} \\ \vdots \\ \bar{y}_{a1} - \bar{y}_{..} \\ \bar{y}_{a2} - \bar{y}_{..} \end{pmatrix} \begin{cases} n_1 \\ \vdots \\ n_a \end{cases}$$

$$(I - P_X) \underline{y} = \begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{a1} \\ y_{a2} \end{pmatrix} \begin{pmatrix} \bar{y}_{11} - \bar{y}_{..} \\ \vdots \\ \bar{y}_{a1} - \bar{y}_{..} \end{pmatrix}$$

$$\textcircled{1} \quad \underline{y}' \underline{y} = \underline{y}' P_W \underline{y} + \underline{y}' (P_X - P_W) \underline{y} + \underline{y}' (I - P_X) \underline{y}$$

Then: (Cochran's theorem)

Let $\underline{y} \sim N(\underline{\mu}, \sigma^2 I)$ and let $A_i, i=1, \dots, K$ be symmetric idempotent matrices with rank s_i .
 If $\sum_{i=1}^K A_i = I$, then $\frac{\underline{y}' A_i \underline{y}}{\sigma^2}$ are independently distributed as $\chi^2(s_i, \phi_i)$, with $\phi_i = \frac{1}{2\sigma^2} \underline{\mu}' A_i \underline{\mu}$ and $\sum_{i=1}^K s_i = n$.

I will now use this result. Note that $P_W, P_X - P_W$ and $I - P_X$ are all idempotent matrices.

$$\begin{aligned} (P_X - P_W)(P_X - P_W) &= P_X^2 - P_W P_X - P_X P_W + P_W^2 \\ &= P_X - P_W - P_W + P_W \end{aligned}$$

~~P_X~~ $\underline{y}' P_W \underline{y}$, $\underline{y}' (P_X - P_W) \underline{y}$ and $\underline{y}' (I - P_X) \underline{y}$ are all independent and they follow χ^2 distributions.

$$\frac{\underline{y}' P_W \underline{y}}{\sim} \sim \chi^2\left(1, \frac{(\underline{x}\beta)' P_W (\underline{x}\beta)}{2}\right)$$

$$\frac{\underline{y}' (P_X - P_W) \underline{y}}{\sim} \sim \chi^2\left(\text{rank}(X) - 1, \frac{(\underline{x}\beta)' (P_X - P_W) \underline{x}\beta}{2}\right)$$

$$\frac{\underline{y}' (\mathbb{I} - P_X) \underline{y}}{\sim} \sim \chi^2(n - \text{rank}(X), \text{# } 0)$$

In the one-way ANOVA example

$$\underline{x}\beta = \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_2 \\ \vdots \\ \mu + \alpha_a \\ \mu + \alpha_1 \end{bmatrix} \quad \left\{ \begin{array}{l} n_1 \\ n_2 \\ \vdots \\ n_a \\ n_1 \end{array} \right\}$$

$$(\underline{x}\beta)' P_W (\underline{x}\beta) = (P_W \underline{x}\beta)' (P_W \underline{x}\beta)$$

$$\text{and } P_W \underline{x}\beta = \frac{1}{(n_1 + \dots + n_a)} J_{n_1 + \dots + n_a} \begin{bmatrix} \mu + \alpha_1 \\ \vdots \\ \mu + \alpha_a \end{bmatrix}$$

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

$$\text{thus } (\underline{x}\beta)' P_W (\underline{x}\beta) = (n_1 + \dots + n_a) \left(\mu + \frac{\sum_{i=1}^a n_i \alpha_i}{\sum_{i=1}^a n_i} \right)^2$$

$$\textcircled{O} (\underline{x}\beta)'(\underline{P}_X - \underline{P}_W)\underline{x}\beta = [(\underline{P}_X - \underline{P}_W)\underline{x}\beta]'[(\underline{P}_X - \underline{P}_W)\underline{x}\beta]$$

$$(\underline{P}_X - \underline{P}_W)\underline{x}\beta = (\underline{P}_X - \underline{P}_W) \begin{bmatrix} \mu + \alpha_1 \\ \vdots \\ \mu + \alpha_a \end{bmatrix}$$

$$\textcircled{O} (\underline{x}\beta)'(\underline{P}_X - \underline{P}_W)(\underline{x}\beta) = \sum_{i=1}^a n_i (\alpha_i - \bar{\alpha})^2 \quad \bar{\alpha} = \textcircled{O} \frac{\sum_{i=1}^a n_i \alpha_i}{\sum_{i=1}^a n_i}$$

$$\frac{\underline{y}' \underline{P}_W \underline{y}}{\sim \chi^2} \sim \chi^2 \left(1, \left(\mu + \sum_{i=1}^a \frac{n_i \alpha_i}{\sum_{i=1}^a n_i} \right)^2 \frac{\sum_{i=1}^a n_i}{2} \right)$$

$$\frac{\underline{y}' (\underline{P}_X - \underline{P}_W) \underline{y}}{\sim \chi^2} \sim \chi^2 \left(\text{rank}(X) - 1, \frac{\sum_{i=1}^a n_i (\alpha_i - \bar{\alpha})^2}{2} \right)$$

$$\frac{\underline{y}' (\underline{I} - \underline{P}_X) \underline{y}}{\sim \chi^2} \sim \chi^2 (n - \text{rank}(X), 0)$$

ANOVA table $n = n_1 + \dots + n_a$

Source	df	Projection	SS	Noncentrality
Mean	1	\underline{P}_W	$\textcircled{O} \underline{y}' \underline{P}_W \underline{y}$	$\frac{1}{2} \left(\mu + \sum_{i=1}^a \frac{n_i \alpha_i}{\sum_{i=1}^a n_i} \right)^2 n$
Group	a-1	$\underline{P}_X - \underline{P}_W$	$\underline{y}' (\underline{P}_X - \underline{P}_W) \underline{y}$	$\frac{1}{2} \sum_{i=1}^a \frac{n_i (\alpha_i - \bar{\alpha})^2}{n_i}$
Error	n-a	$\underline{I} - \underline{P}_X$	$\underline{y}' (\underline{I} - \underline{P}_X) \underline{y}$	0

$$\underline{y}' \underline{P}_W \underline{y} = n \bar{y}_{..}^2$$

$$\underline{y}' (\underline{P}_X - \underline{P}_W) \underline{y} = [(\underline{P}_X - \underline{P}_W) \underline{y}]' [(\underline{P}_X - \underline{P}_W) \underline{y}]$$

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

$$\underline{y}' (\underline{I} - \underline{P}_X) \underline{y} = \sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i..})^2$$

(4)

<u>Source</u>	<u>df</u>	<u>SS</u>	<u>Noncentrality</u>
Mean	1	$n\bar{y}^2$	$\frac{1}{2\pi\nu} \left(\mu + \sum_{i=1}^g n_i \bar{x}_i \right)^2 n$
Group	$a-1$	$\sum_{i=1}^g n_i (\bar{y}_i - \bar{y})^2$	$\frac{1}{2\pi\nu} \sum_{i=1}^g n_i (\bar{x}_i - \bar{x})^2$
Error	$n-a$	$\sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$	0

Recall $\underline{y} \sim N_n(\underline{x}\beta, \sigma^2 I)$

① $\underline{\Delta}^T \hat{\beta} \sim N_s(\underline{\Delta}^T \beta, \sigma^2 \underline{\Delta}^T (\underline{x}^T \underline{x})^{-1} \underline{\Delta})$ where $\underline{\Delta}$ p x s matrix with $\text{rank}(\underline{\Delta}) = s \leq \text{rank}(\underline{x})$.

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

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

④ $\underline{\Delta}^T \hat{\beta}$ has the smallest variance among all linear unbiased estimators of $\underline{\Delta}^T \beta$.

In the normal linear regression the likelihood is

$$f(\underline{y} | \beta, \underline{x}) \propto \frac{1}{(\pi^{\nu})^{n/2}} \exp \left\{ - \frac{(\underline{y} - \underline{x}\beta)^T (\underline{y} - \underline{x}\beta)}{2\sigma^2} \right\}$$

$$\stackrel{Q}{=} \frac{1}{(\pi^{\nu})^{n/2}} \exp \left\{ - \frac{(\underline{y} - \hat{\beta})^T (\underline{y} - \hat{\beta}) - (\beta - \hat{\beta})^T \underline{x}^T \underline{x} (\beta - \hat{\beta})}{2\sigma^2} \right\}$$

~~Since~~ $\hat{\beta}$ = least square estimator, which is also the MLE.

1. ④ sufficient statistic of (β, τ) is

$$(\hat{\beta}, \underbrace{(\underline{y} - \underline{x}\hat{\beta})'(\underline{y} - \underline{x}\hat{\beta})}_{SSE})$$

Testing:

① Testing linear parametric functions, i.e.

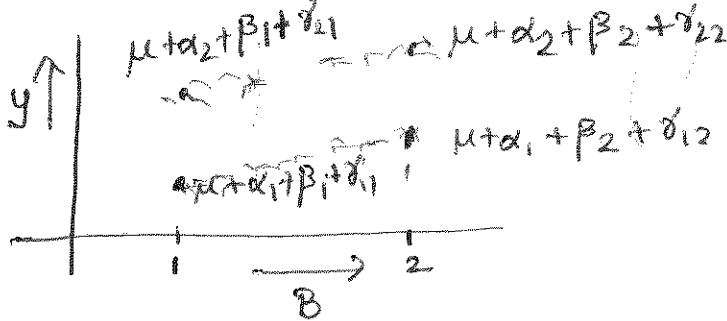
② testing whether $H_0: \underline{\Delta}^T \underline{\beta} = \underline{m}$

② Testing a reduced model.

Examples: $y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijk}$, $i=1 \dots a$,
 $j=1 \dots b$, $k=1 \dots n_{ij}$

$$a=2, b=2$$

$$\text{assume } \delta_{ij}=0 \forall i, j$$

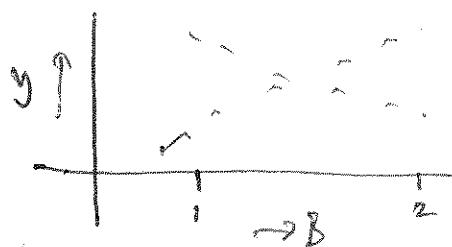


$$y_{11K} = \mu + \alpha_1 + \beta_1 + \delta_{11K}$$

$$y_{12K} = \mu + \alpha_1 + \beta_2 + \delta_{12K}$$

$$y_{21K} = \mu + \alpha_2 + \beta_1 + \delta_{21K}$$

$$y_{22K} = \mu + \alpha_2 + \beta_2 + \delta_{22K}$$



$$H_0: \delta_{11} - \delta_{21} = \delta_{12} - \delta_{22}$$

$$\Leftrightarrow \delta_{11} - \delta_{21} - \delta_{12} + \delta_{22} = 0$$

$$\underline{\beta} = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \delta_{11}, \delta_{12}, \delta_{21}, \delta_{22})$$

$$\Leftrightarrow (0, 0, 0, 0, 0, 1, -1, -1, 1) \underline{\beta} = 0$$

Example: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e$

$$H_0: \beta_2 = 0, \beta_3 = 0 \Leftrightarrow \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \underline{0} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

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
(typically to avoid redundancy)

Also each component of $\underline{\Lambda}^T \underline{\beta}$, given by
 $\underline{\Lambda}_i^T \underline{\beta}$ is estimable.

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

② Recall: $\underline{\Lambda}^T \hat{\underline{\beta}} \sim N(\underline{\Lambda}^T \underline{\beta}, \sigma^2 \underbrace{\underline{\Lambda}^T (\underline{x}' \underline{x})^{-1} \underline{\Lambda}}_{H})$

$$\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m} \sim N(\underline{\Lambda}^T \underline{\beta} - \underline{m}, \sigma^2 H)$$

$$\Rightarrow (\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m})' (\sigma^2 H)^{-1} (\underline{\Lambda}^T \hat{\underline{\beta}} - \underline{m}) \sim \chi^2(s, \phi)$$

$$\phi = \frac{1}{2} (\underline{\Lambda}^T \underline{\beta} - \underline{m})^T (\sigma^2 H)^{-1} (\underline{\Lambda}^T \underline{\beta} - \underline{m})$$

under H_0 , $\phi = 0$

$SSE = \underline{y}' (\mathbf{I} - P_x) \underline{y}$ is independent of $\underline{\Lambda}^T \hat{\underline{\beta}}$.

$$T = \frac{\frac{1}{s} (\underline{\lambda}^T \hat{\beta} - m)^2 + \frac{1}{n-r} (\underline{\lambda}^T \hat{\beta} - m)^2}{\frac{1}{s} \text{SSE} / (n - \text{rank}(X))} \sim \chi^2_{(n-\text{rank}(X), 0)}$$

$$\sim F(s, n - \text{rank}(X), \phi)$$

under $H_0, \phi = 0$

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

we want test of level α ,

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

when $s=1$, that is test $H_0: \underline{\lambda}^T \underline{\beta} = m$

$$\underline{\lambda}^T \hat{\beta} - m \sim N(\underline{\lambda}^T \underline{\beta} - m, \underline{\lambda}^T (\underline{X}' \underline{X})^{-1} \underline{\lambda})$$

$$\Rightarrow \frac{\underline{\lambda}^T \hat{\beta} - m}{\sqrt{\underline{\lambda}^T (\underline{X}' \underline{X})^{-1} \underline{\lambda}}} \sim N(\underline{\lambda}^T \underline{\beta} - m, 1)$$

$$\Rightarrow \frac{\underline{\lambda}^T \hat{\beta} - m}{\sqrt{\frac{1}{s} \text{SSE}} / \sqrt{n - \text{rank}(X)}} \sim t(n - \text{rank}(X), \underline{\lambda}^T \underline{\beta} - m)$$

$$T_2 = \frac{\sqrt{\frac{1}{s} \text{SSE}} / \sqrt{n - \text{rank}(X)}}{\sqrt{\frac{1}{s} \text{SSE}} / \sqrt{n - \text{rank}(X)}}$$

when H_0 is true $T_2 \sim t(n - \text{rank}(X))$

Hence reject H_0 when $|T_2| > t_{n - \text{rank}(X), \frac{\alpha}{2}}$.

back ①