• Result 5.14 If $\mathbf{x}$ is a random variable with $\mathbf{x} \sim N_p(\boldsymbol{\mu}, I)$ and if $\mathbf{M}$ is a perpendicular projection matrix, then $\mathbf{x}^T\mathbf{M}\mathbf{x} \sim \chi^2(r(\mathbf{M}), \boldsymbol{\mu}^T\mathbf{M}\boldsymbol{\mu}/2)$.

$y^T P y$

$y^T y = y^T I y$

$y^T (I-P) y$

$\mu \in \mathcal{C}(M)$

$\Rightarrow M \cdot \mu = \mu$

$\frac{\mu^T M \mu}{2} = \frac{(M\mu)^T (M\mu)}{2}$

$= \frac{\mu^T \mu}{2}$

√ Observe if $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \sigma^2 I)$, then $\mathbf{x}/\sigma \sim N_p((1/\sigma)\boldsymbol{\mu}, I)$ and $\mathbf{x}^T\mathbf{M}\mathbf{x}/\sigma^2 \sim \chi^2(r(\mathbf{M}), \boldsymbol{\mu}^T\mathbf{M}\boldsymbol{\mu}/(2\sigma^2))$.

$I, \sigma^2 I$

- Result 5.15 Let $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$ with $\mathbf{V}$ nonsingular, and let $\mathbf{A}$ be a symmetric matrix; then if $\underline{\mathbf{AV}}$ is idemptent with rank s, then $\mathbf{x}^T\mathbf{A}\mathbf{x} \sim \chi^2(s, \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}/2)$.

~~Result 5.14~~

- ~~Lemma: If $\mathbf{x} \sim N_n(\boldsymbol{\mu}, I)$, where $\boldsymbol{\mu} \in C(\mathbf{M})$ and if $\mathbf{M}$ is a perpendicular projection matrix, then $\mathbf{x}^T\mathbf{x} \sim \chi^2(r(\mathbf{M}), \boldsymbol{\mu}^T\boldsymbol{\mu}/2)$.~~

$\frac{\mu^T M \mu}{2}$

- Result 5.16 Let $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$ and let $\mathbf{A}$ be a symmetric matrix with rank s; then if $\mathbf{BVA} = \mathbf{0}$, then $\mathbf{Bx}$ and $\mathbf{x}^T\mathbf{Ax}$ are independent.

$\hat{y} = Py \qquad y^T(I-P)y = \hat{e}^T\hat{e}$

- Cor 5.4 Let $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$, $\mathbf{A}$ be a symmetric matrix with rank r, and $\mathbf{B}$ be symmetric with rank s; if $\mathbf{BVA} = \mathbf{0}$, then $\mathbf{x}^T\mathbf{Ax}$ and $\mathbf{x}^T\mathbf{Bx}$ are independent.

$$\hat{y}^T\hat{y} = y^T P y = (\underbrace{Py}_{\hat{y}})^T(\underbrace{Py}_{\hat{y}}) = y^T \underbrace{P^T P}_{= PP = P} y \qquad P^T = P, \quad PP = P$$

$$\hat{e}^T\hat{e} = y^T(I-P)y$$

$r(P) = \dim(\mathcal{C}(X)) = r(X) = r$

$$y \sim N_n(\overset{n\times p}{X}\beta, \sigma^2 I)$$

- Let's connect this to what we have $\Leftrightarrow \frac{y}{\sigma} \sim N_n\left(\frac{X\beta}{\sigma}, I\right)$

$$\mathbf{y}^T\mathbf{y} = \hat{\mathbf{y}}^T\hat{\mathbf{y}} + \hat{\mathbf{e}}^T\hat{\mathbf{e}} = \mathbf{y}^T\mathbf{P}\mathbf{y} + \mathbf{y}^T(I - \mathbf{P})\mathbf{y}.$$

where $\mathbf{P}$ is the orthogonal projection operator onto $C(\mathbf{X})$.

√ What is the distribution of $\mathbf{y}^T\mathbf{y}$?

$\left(\frac{y}{\sigma}\right)^T \underset{n\times n}{I} \left(\frac{y}{\sigma}\right) \qquad \frac{y^Ty}{\sigma^2} \sim \chi^2\left(n, \frac{1}{2}\left(\frac{X\beta}{\sigma}\right)^T I \left(\frac{X\beta}{\sigma}\right)\right)$

$\frac{y^Ty}{\sigma^2} \sim \chi^2\left(n, \frac{\beta^TX^TX\beta}{2\sigma^2}\right)$

$y^Ty \sim \sigma^2 \chi^2\left(n, \frac{\beta^TX^TX\beta}{2\sigma^2}\right)$

√ What is the distribution of $\hat{\mathbf{y}}^T\hat{\mathbf{y}}$?

$\left(\frac{y}{\sigma}\right)^T P \left(\frac{y}{\sigma}\right) \sim \chi^2\left(r, \frac{\beta^TX^TPX\beta}{2\sigma^2}\right)$

$= \frac{\beta^TX^TX\beta}{2\sigma^2}$

√ What is the distribution of $\hat{\mathbf{e}}^T\hat{\mathbf{e}}$?

$\left(\frac{y^T}{\sigma}\right)(I-P)\left(\frac{y}{\sigma}\right) \sim \chi^2\left(n-r, \underbrace{\frac{\beta^TX^T(I-P)X\beta}{2\sigma^2}}_{=0}\right)$

$$y \sim N_n(0, \underbrace{\sigma^2 I}_{=V}) \qquad \hat{y} = \overset{=B}{P}y \qquad \begin{aligned} &BVA \\ &= P\sigma^2 I \cdot (I-P) \\ &= \sigma^2(P-P^2) \\ &= 0 \end{aligned}$$

$$\hat{e}^T\hat{e} = y^T\underbrace{(I-P)}_{=A}y$$

- Let's connect this to what we have (contd)

$$\mathbf{y}^T\mathbf{y} = \hat{\mathbf{y}}^T\hat{\mathbf{y}} + \hat{\mathbf{e}}^T\hat{\mathbf{e}} = \mathbf{y}^T\mathbf{P}\mathbf{y} + \mathbf{y}^T(I - \mathbf{P})\mathbf{y}.$$

where $\mathbf{P}$ is the orthogonal projection operator onto $C(\mathbf{X})$.

✓ What is the distribution of $\hat{\sigma}^2$?

$\hat{\sigma}^2 = \frac{SSE}{n-r} = MSE.$ $\quad SSE = \hat{e}^T\hat{e}$

$$\frac{\hat{e}^T\hat{e}}{\sigma^2} \sim \chi^2(n-r)$$

✓ Are $\hat{\mathbf{y}}^T\hat{\mathbf{y}}$ and $\hat{\mathbf{e}}^T\hat{\mathbf{e}}$ independent?

$= y^T\underbrace{P}_{=A}y \qquad = y^T\underbrace{(I-P)}_{=B}y$

$y \sim N_n(X\beta, \underbrace{\sigma^2 I}_{=V})$

$$BVA = (I-P)\sigma^2 I P = \sigma^2(I-P)P = \sigma^2(P-P^2) = \sigma^2 \times 0$$

✓ What is the distribution of $\frac{||\hat{\mathbf{y}}||^2/r}{||\hat{\mathbf{e}}||^2/(n-r)}$?

$$\frac{||\hat{y}||^2/r}{||\hat{e}||^2/(n-r)} = \frac{\frac{\hat{y}^T\hat{y}}{\sigma^2}/r}{\frac{\hat{e}^T\hat{e}}{\sigma^2}/(n-r)} \sim F\left(r, n-r, \frac{\beta^T X^T X\beta}{2\sigma^2}\right)$$

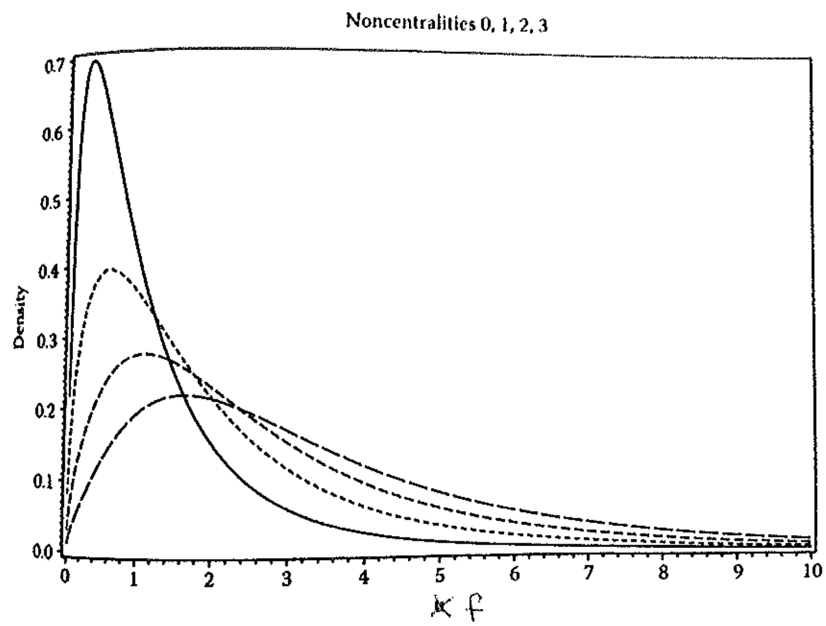
† Def 5.7: Let u_1 and u_2 be independent random variables, with $u_1 \sim \chi^2(p_1)$ and $u_2 \sim \chi^2(p_2)$; then $f \equiv \frac{u_1/p_1}{u_2/p_2}$ has the F-distribution with p_1 and p_2 degrees of freedom, denoted as $f \sim F(p_1, p_2)$.

† Def 5.8: 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 $f \equiv \frac{u_1/p_1}{u_2/p_2}$ has the F-distribution with p_1 and p_2 degrees of freedom, noncentrality ϕ, denoted as $f \sim F(p_1, p_2, \phi)$.

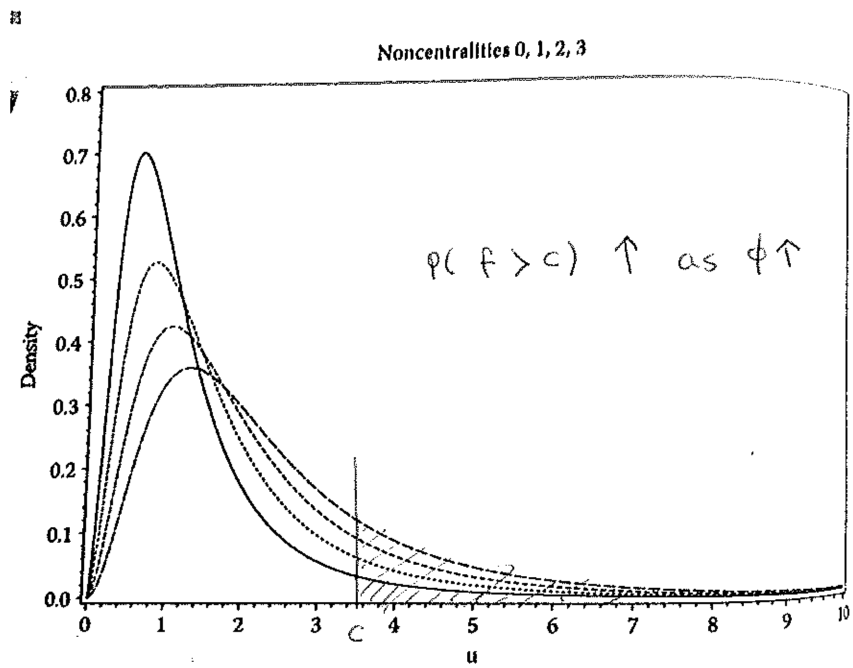
$$E(f) = \frac{p_2(p_1+\phi)}{p_1(p_2-1)}, \quad p_2 > 2$$

· For fixed p_1 & p_2, $E(F) \uparrow$ as $\phi \uparrow$

• Figure 5.3: Noncentral F densities with $p_1 = 3$ and $p_2 = 10$ degrees of freedom. Increasing ϕ flattens the density and shifts the distribution to the right



- Figure 5.3: Noncentral F densities with $p_1 = 6$ and $p_2 = 10$ degrees of freedom. Increasing ϕ flattens the density and shifts the distribution to the right



- Result 5.13 Let $w \sim F(p_1, p_2, \phi)$, then for fixed p_1 and p_2 and $c > 0$ $P(w > c)$ is strictly increasing in ϕ.

† Def 5.9: Let $u \sim \mathrm{N}(\mu, 1)$ and $v \sim \chi^2(k)$. If u and v are independent, then $t = u/\sqrt{v/k}$ has the noncentered Student's t-distribution with k degrees of freedom and noncentrality μ, denoted by $t \sim t(k, \mu)$. If $\mu = 0$, the distribution is generally known as Student's t, denoted by $t \sim t(k)$.

√ If $t \sim t(k, \mu)$, then $t^2 \sim F(1, k, \mu^2/2)$.

$$E(t) = \mu \cdot \sqrt{\frac{k}{2}} \cdot \frac{\Gamma\left(\frac{k-1}{2}\right)}{\Gamma(k/2)}, \quad k > 1$$

- Let's connect this to what we have

$$\mathbf{y} \sim \mathrm{N}_n(\mathbf{X}\beta, \sigma^2 I)$$

√ What is the distribution of $\frac{(\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}} - \boldsymbol{\lambda}^T\boldsymbol{\beta})}{\sqrt{MSE\boldsymbol{\lambda}^T(\mathbf{X}^T\mathbf{X})^{-}\boldsymbol{\lambda}}}$ for $\boldsymbol{\lambda}^T\boldsymbol{\beta}$ estimable?

From the previous slide, $\lambda^T\hat{\beta} \sim N(\lambda^T\beta, \sigma^2\lambda^T(X^TX)^{-}\lambda)$

$\cancel{|\lambda^T\hat{\beta} - \lambda^T\beta|}$ $\quad \frac{SSE}{\sigma^2} \sim \chi^2(n-r, 0)$

$$\Rightarrow \frac{(\lambda^T\hat{\beta} - \lambda^T\beta)}{\sqrt{\cancel{\sigma^2}\lambda^T(X^TX)^{-}\lambda}} \Big/ \sqrt{\frac{SSE}{\cancel{\sigma^2}}/(n-r)} \qquad \frac{SSE}{n-r} = MSE$$

$$= \frac{(\lambda^T\hat{\beta} - \lambda^T\beta)}{\sqrt{MSE\,\lambda^T(X^TX)^{-}\lambda}} \sim t(n-r)$$

$$\frac{\lambda^T\hat{\beta}}{\sqrt{MSE\,\lambda^T(X^TX)^{-}\lambda}} \sim t(n-r, \lambda^T\beta)$$

$y_{ij} = \mu + \alpha_i + e_{ij}$

$i = 1, 2, 3, 4$

$j = 1, \cdots, n_i$

$(4, 6, 6, 8)$

$\boxed{n = 24}$

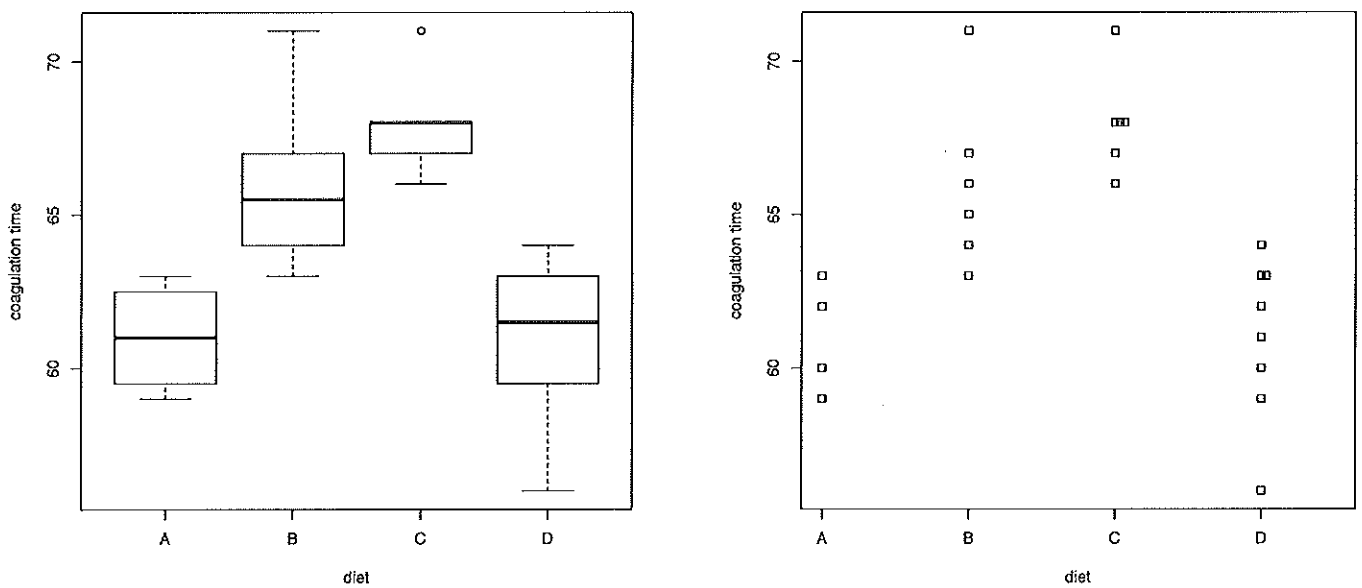
Coagulation Example– revisit

Dataset comes from a study of blood coagulation times. 24 animals were randomly assigned to four different diets and the samples were taken in a random order (taken from Linear Models with R, page 182)

4

```
> rm(list=ls(all=TRUE))
> library(faraway)
> data(coagulation)
>
> plot(coag~diet, coagulation, ylab="coagulation time")
> with(coagulation, stripchart(coag ~ diet, vertical=TRUE, metho
```

Coagulation Example– revisit



♠ ex One-way ANOVA – Coagulation Example

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

Find the decomposition of sums of squares.

Coagulation Example– revisit

```
> options(contrasts=c("contr.sum", "contr.poly"))
> g2 <- lm(coag ~ diet, coagulation)
> summary(g2)
Call:
lm(formula = coag ~ diet, data = coagulation)
Residuals:
   Min     1Q Median     3Q    Max
 -5.00  -1.25   0.00   1.25   5.00
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  64.0000     0.4979 128.537  < 2e-16 ***
diet1        -3.0000     0.9736  -3.081 0.005889 **
diet2         2.0000     0.8453   2.366 0.028195 *
diet3         4.0000     0.8453   4.732 0.000128 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
Residual standard error: 2.366 on 20 degrees of freedom
Multiple R-squared:  0.6706, Adjusted R-squared:  0.6212
F-statistic: 13.57 on 3 and 20 DF,  p-value: 4.658e-05
```

Coagulation Example– revisit

```
> anova(g2)
Analysis of Variance Table
                                  SS/df
Response: coag
          Df Sum Sq Mean Sq F value    Pr(>F)
diet       3    228    76.0  13.571 4.658e-05 ***
Residuals 20    112     5.6
---              340
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

★ More common ANOVA table

Source	Deg. of Freedom	Sum of Squares	Mean Square	F
Model (err) Regression	$\overset{r}{p} - 1$	SS_{reg} SSerr SSM	$SS_{reg}/(\overset{r}{\cancel{p}} - 1)$	F
Residual	$n - p^{r}$	RSS	$RSS/(n - \overset{r}{\cancel{p}})$	
Total	$n - 1$	TSS		

Table 3.1 *Analysis of variance table.*

- **ANOVA table**

Source	df	Projection	SS	Noncentrality
Mean	1	$\mathbf{P}_1$	$SSM = n\bar{y}^2$	$\frac{1}{2}n(\mu + \bar{\alpha})^2/\sigma^2$
Group	a-1	$\mathbf{P}_x - \mathbf{P}_1$	$SSA = \sum_{i=1}^{a} n_i \bar{y}_{i\cdot}^2 - n\bar{y}^2$	$\frac{1}{2}\sum_{i=1}^{a}(\alpha_i - \bar{\alpha})^2/\sigma^2$
Error	n-a	$I - \mathbf{P}_x$	$SSE = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2$	0

• Th 5.1 (Cochran's Theorem) Let $\mathbf{y} \sim \mathrm{N}_n(\overset{X\beta}{\overset{\|}{\boldsymbol{\mu}}}, \sigma^2 I)$, and let $\mathbf{A}_i$, $i = 1, \ldots, k$ be symmetric idempotent matrices with rank s_i. If $\sum_{i=1}^k \mathbf{A}_i = I$, then $(1/\sigma^2)\mathbf{y}^T \mathbf{A}_i \mathbf{y}$ are independently distributed as $\chi^2(s_i, \phi_i)$, with $\phi_i = \frac{1}{2\sigma^2}\boldsymbol{\mu}^T \mathbf{A}_i \boldsymbol{\mu}$ and $\sum_{i=1}^k r_i = n$.

Coagulation

$$y_{ij} = \mu + \alpha_i + e_{ij} \qquad i = 1, 2, 3, 4 \qquad j = 1, \dots, n_i \; (4, 6, 6, 8)$$

$n = 24$

$$\underbrace{\begin{bmatrix} y_{11} \\ \vdots \\ y_{14} \\ y_{21} \\ \vdots \\ \vdots \\ y_{48} \end{bmatrix}}_{24\times 1} = \underbrace{\begin{bmatrix} 1_4 & 1_4 & 0_4 & 0_4 & 0_4 \\ 1_6 & 0_6 & 1_6 & 0_6 & 0_6 \\ 1_6 & 0_6 & 0_6 & 1_6 & 0_6 \\ 1_8 & 0_8 & 0_8 & 0_8 & 1_8 \end{bmatrix}}_{24\times 5} \underbrace{\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}}_{5\times 1} + \begin{bmatrix} e_{11} \\ \vdots \\ \vdots \\ e_{48} \end{bmatrix}$$

W = first column, X = whole matrix

$P_X, \; P_W = (P_1)$

$\Rightarrow$ observe $\mathcal{C}(W) \subseteq \mathcal{C}(X)$

\# From our previous lecture,

If $\mathcal{C}(W) \subseteq \mathcal{C}(X)$, $P_X - P_W$ is the projection onto $\mathcal{C}((I - P_W)X)$

$$P_1 = P_W = 1(1^T 1)^{-1} 1^T = \frac{1}{24} 1 1^T = \frac{1}{24} J_{24\times 24} \qquad \dim(\mathcal{C}(W)) = 1$$

$$\frac{1}{24}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_{11} \\ \vdots \\ y_{48} \end{bmatrix}$$

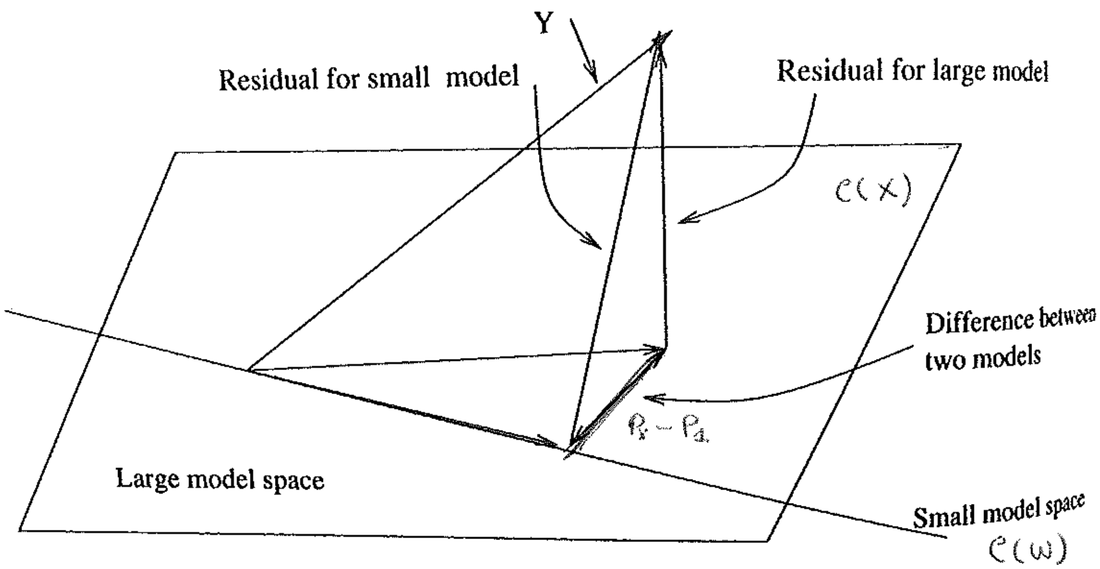
$$P_1 Y = \begin{bmatrix} \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{\cdot\cdot} \end{bmatrix}_{24\times 1}$$

$$\begin{array}{l} \tfrac14 \; \tfrac14 \; \tfrac14 \; \tfrac14 \; 0 \cdots\cdots 0 \\ P_X = X(X^TX)^{-}X^T = \begin{bmatrix} \boxed{\frac{1}{4} J_{4\times 4}} & O_{4\times 6} & O_{4\times 6} & O_{4\times 8} \\ O_{6\times 4} & \boxed{\frac{1}{6} J_{6\times 6}} & O_{6\times 6} & O_{6\times 8} \\ O_{6\times 4} & O_{6\times 6} & \boxed{\frac{1}{6} J_{6\times 6}} & O_{6\times 8} \\ O_{8\times 4} & O_{8\times 6} & O_{8\times 6} & \boxed{\frac{1}{8} J_{8\times 8}} \end{bmatrix} \begin{bmatrix} y_{11} \\ \vdots \\ y_{14} \\ y_{21} \\ \vdots \\ y_{48} \end{bmatrix} \end{array}$$

$\dim(\mathcal{C}(X)) = 4 \uparrow 24\times 5$

$$P_X y = \begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{1\cdot} \\ \vdots \end{bmatrix}$$

★ Figures help!



1.5

$$P_X\, y = \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \vdots \\ \bar{y}_{4\cdot} \end{bmatrix}$$

$y_{ij} = \mu + e_{ij} \Rightarrow P_1 y = \hat{y}_A$

$y_{ij} = \mu + \alpha_i + e_{ij} \Rightarrow P_X y = \hat{y}_B$

$\| \hat{y}_A - \hat{y}_B \|^2$

$C(W) \subset C(X)$

$$y^T(P_X - P_1)y$$
$$\parallel$$
$$y^T P_1 y + y^T P_1 y + y^T P_X y + y^T(I - P_X)y$$

$$\forall \quad y^T y = y^T P_X y + y^T(I - P_X)y$$

$$\underbrace{y^T y}_{} = \underbrace{y^T P_1 y}_{} + \underbrace{y^T(P_X - P_1)y}_{} + \underbrace{y^T(I - P_X)y}_{}$$

→ projection onto $C((I - P_1)X)$

$C(P_X - P_1) = C(W)^{\perp}_{C(X)}$

projection onto the orthogonal complement of $C(W)$ with respect to $C(X)$

Q : the distribution of the parts ?

$$y^T y - y^T P_1 y = y^T(I - P_1)y \; : \; SST$$

$$y^T(I - P_1)y = (y - P_1 y)^T(y - P_1 y) = [\, y_{11} - \bar{y}_{\cdot\cdot}, \;\ldots\;, y_{48} - \bar{y}_{\cdot\cdot} \,] \begin{bmatrix} y_{11} - \bar{y}_{\cdot\cdot} \\ \vdots \\ y_{48} - \bar{y}_{\cdot\cdot} \end{bmatrix} = \sum_i \sum_j (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

$$y^T(P_X - P_1)y = (P_X y - P_1 y)^T (P_X y - P_1 y)$$

$$= [\, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \;\ldots\;, \bar{y}_{4\cdot} - \bar{y}_{\cdot\cdot} \,] \begin{bmatrix} \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot} \\ \vdots \\ \bar{y}_{4\cdot} - \bar{y}_{\cdot\cdot} \end{bmatrix} = \sum_i \sum_j (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

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

$$y^T(I - P_X)y = (y - P_X y)^T(y - P_X y) = [\, y_{11} - \bar{y}_{1\cdot}, \;\ldots\;, y_{48} - \bar{y}_{4\cdot} \,] \begin{bmatrix} y_{11} - \bar{y}_{1\cdot} \\ \vdots \\ y_{48} - \bar{y}_{4\cdot} \end{bmatrix}$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2$$

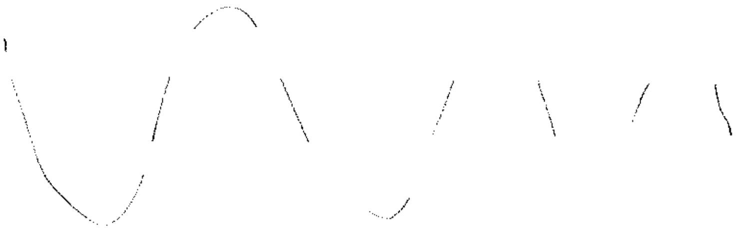
$$y^T y - y^T P_1 y = y^T (P_x - P_1) y + y^T (I - P_x) y$$

$$\underbrace{y^T (I - P_1) y}_{= SST} = \underbrace{y^T (P_x - P_1) y}_{\substack{= SSM \\ = SS\text{-}trt \\ = SSreg}} + \underbrace{y^T (I - P_x) y}_{\substack{= RSS \\ = SSE}}$$

P_1 : $\underbrace{\mu + \alpha_1}_{\alpha_2 = \alpha_3 = 0}$

P_X :

2.5



$u \sim \chi^2(r, \phi) \quad E(u) = r + 2\phi$

$$\Rightarrow \quad y^T y = y^T P_1 y + y^T (P_x - P_1) y + y^T (I - P_x) y$$

$$\frac{y^T P_1 y}{\sigma^2} \sim \chi^2\Big(\underbrace{r(P_1)}_{=1}, \overbrace{\frac{(X\beta)^T P_1 (X\beta)}{2\sigma^2}}^{= \frac{24(\mu + \bar{\alpha})^2}{2\sigma^2}}\Big) \quad \approx \bar{y}_{..}$$

$$(X\beta)^T P_1 X\beta = (P_1 X\beta)^T (P_1 X\beta) = \left(\mu + \underbrace{\frac{\sum n_i \alpha_i}{24}}_{\equiv \bar{\alpha}}\right)^2 \cdot 24$$

$$= 24(\mu + \bar{\alpha})^2$$

$\dim(C(X))$
$= r(P_x) - r(P_1)$
$= a - 1 = 4 - 1 = 3$

$$\frac{y^T (P_x - P_1) y}{\sigma^2} \sim \chi^2\left(a-1, \frac{(X\beta)^T (P_x - P_1) X\beta}{2\sigma^2}\right)$$

$$= \frac{\sum_{i=1}^{4} n(\alpha_i - \bar{\alpha})^2}{2\sigma^2}$$

$$\frac{y^T (I - P_x) y}{\sigma^2} \sim \chi^2\Big(\underbrace{n - r}_{= 24 - 4 = 20}, \underbrace{\frac{(X\beta)^T (I - P_x) X\beta}{2\sigma^2}}_{= 0}\Big)$$

	SS	df
Grand Mean	$y^T P_1 y = N\bar{y}_{..}$	$1 = \dim(P_1)$
Treatments	$y^T(P_X - P_1)y = \sum_i n_i(\bar{y}_{i.} - \bar{y}_{..})^2$	$a-1 = \dim(P_X - P_1)$
Error	$y^T(I - P_X)y = \sum_i\sum_j(y_{ij} - \bar{y}_{i.})^2$	$n-a = \dim(I - P_X)$
Total	$y^T y$	$n = \dim(I)$

$\boxed{E(SS)}$

$$\sigma^2\left(1 + 2\times\frac{N(\mu+\bar{\alpha})^2}{2\sigma^2}\right) = \sigma^2 + N(\mu+\bar{\alpha})^2$$

$$\sigma^2\left(a-1 + 2\times\frac{\sum_{i=1}^{a} n_i(\alpha_i-\bar{\alpha})^2}{2\sigma^2}\right) = \sigma^2(a-1) + \sum_{i=1}^{n} n_i(\alpha_i-\bar{\alpha})^2$$

$$\sigma^2(n-r)$$

$MS = SS/df$ $\Rightarrow$ $E(MS)$

$$\sigma^2 + N(\mu+\bar{\alpha})^2$$

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

$$\sigma^2$$

Ronald book p 41 Table 4.1

AMS 256
Monahan Chapter 6: Statistical Inference & Ronald Chapter 3: Testing Hypotheses

Spring 2016

† The form of linear models is

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e},$$

- $\mathbf{y}$: $n \times 1$ vector of observations (random)
- $\mathbf{X}$: $n \times p$ matrix of known constants (*design* matrix) with $r(\mathbf{X}) = r$
- β: $p \times 1$ vector of unobservable parameters
- $\mathbf{e}$: $n \times 1$ vector of unobservable <u>random</u> errors

♠ Assume $\mathbf{e} \sim \mathrm{N}_n(\mathbf{0}, \sigma^2 I)$.

- Recall

$X: n \times p, \quad r(X) = r$

  - $\mathbf{y} \sim \mathrm{N}_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I)$
  - $\Lambda^T \hat{\boldsymbol{\beta}} \sim \mathrm{N}_s(\Lambda^T \boldsymbol{\beta}, \sigma^2 \Lambda^T (\mathbf{X}^T\mathbf{X})^- \Lambda)$ where Λ: $p \times s$ matrix with $rank(\Lambda) = s \le rank(\mathbf{X})$.
  - $SSE/\sigma^2 = \mathbf{y}^T(I - \mathbf{P})\mathbf{y}/\sigma^2 \sim \chi^2(n-r)$.

    $\hat{e}^T\hat{e}/\sigma^2$

    $a^T \in \mathcal{C}(X^T)$
  - $\Lambda^T \hat{\boldsymbol{\beta}}$ and SSE/σ^2 are independent.

$\Lambda^T\hat{\beta} = \Lambda^T (X^TX)^- X^Ty = a^T X(X^TX)^- X^Ty = a^T\hat{y}$

  - $\Lambda^T \hat{\boldsymbol{\beta}}$ has the smallest variance among all ~~**linear**~~, unbiased estimators (BLUE).

★ The likelihood is:

$$
\begin{aligned}
f(\mathbf{y} \mid \beta, \sigma^2) &= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta)\right\} \\
&= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y}^T\mathbf{y}^T + 2\mathbf{y}^T\mathbf{X}\beta + (\mathbf{X}\beta)^T(\mathbf{X}\beta)\right\}
\end{aligned}
$$

★ Observe that β, σ^2

- $(\mathbf{y}^T\mathbf{y}, \mathbf{X}^T\mathbf{y})$ are a minimal sufficient statistic (equivalently, $(SSE, \mathbf{X}^T\mathbf{y})$, Result 6.1 and Cor 6.1)
- $(\mathbf{y}^T\mathbf{y}, \mathbf{X}^T\mathbf{y})$ are a complete sufficient statistic (Result 6.2)

• Result 6.3 In the normal linear model $\mathbf{y} \sim N_n(\mathbf{X}\beta, \sigma^2 I)$ with unknown parameters (β, σ^2), $(\hat{\beta}, SSE/n)$ is a maximum likelihood estimator of (β, σ^2) where $\hat{\beta}$ solves the normal equations and $SSE = \mathbf{y}^T(I - \mathbf{P})\mathbf{y}$.

MLE $\hat{\sigma}^2 = \frac{SSE}{n}$ unbiased $\hat{\sigma}^2 = \frac{SSE}{n-r}$

$\Rightarrow$ Check!!! HW

• Cor 6.3 Under the normal linear model, the maximum likelihood estimator of an estimable function $\Lambda^T\beta$ is $\Lambda^T\hat{\beta}$, where $\hat{\beta}$ solves the normal equations. invariance property of MLE

* Observe $\Lambda^T\hat{\beta} = \Lambda^T(\mathbf{X}^T\mathbf{X})^{-}\mathbf{X}^T\mathbf{y}$ ∘ Rao - Blackwell

• Cor 6.2 Under the normal Gauss-Markov model, the LSE $\Lambda\hat{\beta}$ of an estimable function $\Lambda^T\beta$ has the smallest variance among *all unbiased estimator.* ~~linear~~

⇔ In summary, with normal errors, LSEs are best estimators among all unbiased estimates.

• The $MSE = SSE/(n-r)$ is a minimum variance unbiased estimate of σ^2.