Stat 565
1-11-18

Cochran's Theorem

①

Let $Z_1, Z_2, \ldots, Z_v$ be iid $N(0,1)$

Let $\sum_{i=1}^{v} Z_i^2 = Q_1 + Q_2 + \ldots + Q_k$

where $Q_i \sim \chi^2_{v_i}$

Then the Q_i's are independent iff

$v_1 + v_2 + \ldots + v_k = v$.

②

From last time, $SS_T = SS_{TRT} + SS_E$

$$\begin{array}{ccc} SS_T & SS_{TRT} & SS_E \\ \downarrow & \downarrow & \downarrow \\ N-1 & a-1 & N-a \end{array}$$

If we assume that the ε_{ij}'s are normally distributed, then under H_0,

SS_T, SS_{TRT}, and SS_E will all have χ^2 distributions

$\therefore$ By Cochran's Theorem, SS_{TRT} & SS_E are independent.

So $$\frac{\frac{SS_{TRT}}{df_{TRT}}}{\frac{SS_E}{df_E}} \sim F_{df_{TRT},\, df_E}$$

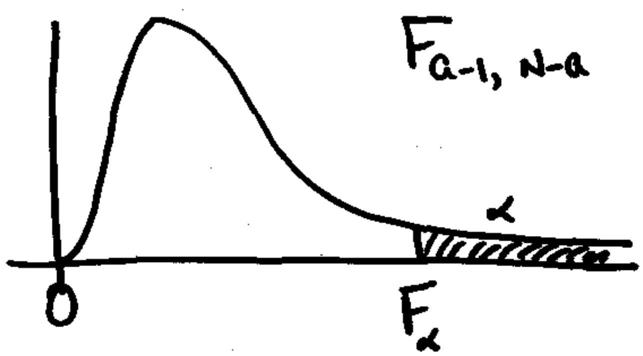
ANOVA table

Source	SS	df	MS	F
TRT	SS_{TRT}	$a-1$	$SS_{TRT}/a-1$	MS_{TRT}/MS_E
ERR	SS_E	$N-a$	$SS_E/N-a$	
TOT	SS_T	$N-1$	—	

4

$H_0: \tau_1 = \tau_2 = \ldots = \tau_a = 0$

H_1: The τ_i's are not all 0



Reject H_0 if $F > F_\alpha$

Parameter estimation

$i = 1, \ldots, a$

$j = 1, \ldots, n$

Model: $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ $\qquad \sum_{i=1}^{a} \tau_i = 0$

$\mu_i = \mu + \tau_i$

(5)

1st method: Use least squares

$$SS_E = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \mu - \tau_i)^2$$

$$\frac{\partial SS_E}{\partial \mu} = \sum_i \sum_j \cancel{2}(y_{ij} - \mu - \tau_i)(-\cancel{1}) \stackrel{set}{=} 0$$

$$y_{..} - N\mu - 0 = 0$$

$$\hat{\mu} = \frac{y_{..}}{N} = \bar{y}_{..}$$

(6)

$$\frac{\partial SS_E}{\partial \tau_i} = \sum_{j=1}^{n} \cancel{2}(y_{ij} - \mu - \tau_i)(-\cancel{1}) \stackrel{set}{=} 0$$

$$y_{i.} - n\mu - n\tau_i = 0$$

$$\hat{\tau}_i = \frac{y_{i.}}{n} - \hat{\mu}$$

$$= \bar{y}_{i.} - \bar{y}_{..}$$

2nd method: Use the matrix version

⑦

$$y_{ij} = \mu_i + \varepsilon_{ij} \qquad \begin{matrix} i=1,\dots,a \\ j=1,\dots,n \end{matrix}$$

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{1n} \\ \hline y_{21} \\ \vdots \\ y_{2n} \\ \hline \vdots \\ \vdots \\ \hline y_{a1} \\ \vdots \\ y_{an} \end{bmatrix}_{N\times 1} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \\ \hline 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & 0 & \cdots & 0 \\ \hline & & \vdots & & \\ \hline 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}_{N\times a} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_a \end{bmatrix}_{a\times 1} + \vec{\varepsilon}_{N\times 1}$$

$$Y = X\beta + \vec{\varepsilon}$$

⑧

From linear regression, we know

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

$$\underset{a\times a}{X'X} = \underbrace{\begin{bmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ & & & \\ & & & 1 \cdots 1 \end{bmatrix}_{a\times N}}_{X'} \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & & & \\ 1 & 0 & \cdots & 0 \\ \hline & \vdots & & \\ \hline 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \end{bmatrix}_{N\times a}}_{X}$$

$$= \begin{bmatrix} n & & 0 \\ & \ddots & \\ 0 & & n \end{bmatrix}_{a\times a} = n\, I_{a\times a}$$

9

$$(X'X)^{-1} = \frac{1}{n} I_{a\times a}$$

$$\underset{a\times 1}{X'Y} = \underset{a\times N}{\begin{bmatrix} 1 \cdots 1 & & 0 \cdots 0 \\ 0 \quad 0 & & \vdots \quad \vdots \\ \vdots \quad \vdots & \cdots & \vdots \quad \vdots \\ 0 \quad 0 & & 0 \quad 0 \\ 0 \quad 0 & & 1 \cdots 1 \end{bmatrix}} \underset{N\times 1}{\begin{bmatrix} y_{11} \\ \vdots \\ y_{1n} \\ \hline \vdots \\ \hline y_{a1} \\ \vdots \\ y_{an} \end{bmatrix}}$$

$$= \begin{bmatrix} y_{1\cdot} \\ y_{2\cdot} \\ \vdots \\ y_{a\cdot} \end{bmatrix}$$

Now $\hat{\beta} = (X'X)^{-1}X'Y$

$$= \frac{1}{n} I \begin{bmatrix} y_{1\cdot} \\ \vdots \\ y_{a\cdot} \end{bmatrix} = \begin{bmatrix} \bar{y}_{1\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix}$$

10

That is, $\hat{\mu}_i = \bar{y}_{i\cdot}$

This is consistent with our least squares solution of $\hat{\mu} = \bar{y}_{\cdot\cdot}$ and $\hat{\tau}_i = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}$

Confidence intervals for individual parameters

(Assume $\varepsilon_{ij} \sim$ iid $N(0, \sigma^2)$)

Then $y_{ij} \sim$ indep $N(\mu + \tau_i, \sigma^2)$

11

So $\bar{y}_{i\cdot} \sim N(\underbrace{\mu + \tau_i}_{\mu_i}, \sigma^2/n)$

And $\dfrac{\bar{y}_{i\cdot} - \mu_i}{\sigma/\sqrt{n}} \sim N(0,1)$

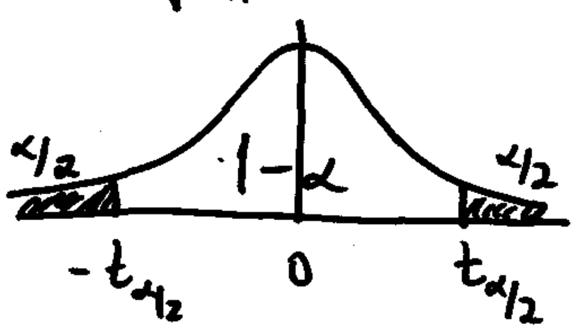
indep.

Also, $\dfrac{SS_E}{\sigma^2} \sim \chi^2_{N-a}$

So $$\frac{\frac{\bar{y}_{i\cdot} - \mu_i}{\cancel{\sigma}/\sqrt{n}}}{\sqrt{\frac{SS_E}{\cancel{\sigma^2}}/(N-a)}} \sim t_{N-a}$$

12

$$\frac{\bar{y}_{i\cdot} - \mu_i}{\sqrt{\frac{MS_E}{n}}} \sim t_{N-a}$$



$$P\left(-t_{\alpha/2} < \frac{\bar{y}_{i\cdot} - \mu_i}{\sqrt{\frac{MS_E}{n}}} < t_{\alpha/2}\right) = 1 - \alpha$$

$$\bar{y}_{i.} - t_{\alpha/2}\sqrt{\frac{MS_E}{n}} < \mu_i < \bar{y}_{i.} + t_{\alpha/2}\sqrt{\frac{MS_E}{n}}$$

is a $(1-\alpha)100\%$ confidence interval

for μ_i

HW #1 due Thurs Jan 18

p.130 3.7(a)

3.12(a)

3.7. The tensile strength of Portland cement is being studied. Four different mixing techniques can be used economically. A completely randomized experiment was conducted and the following data were collected:

Mixing Technique	**Tensile Strength (lb/in^2)**			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

(a) Test the hypothesis that mixing techniques affect the strength of the cement. Use $\alpha = 0.05$.

3.12. A pharmaceutical manufacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

Dosage	**Observations**			
20 g	24	28	37	30
30 g	37	44	31	35
40 g	42	47	52	38

(a) Is there evidence to indicate that dosage level affects bioactivity? Use $\alpha = 0.05$.