Stat 565

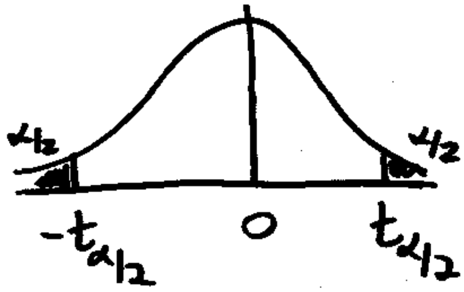
1-16-18

1

Hypothesis tests for parameters in ANOVA

$H_0: \mu_i = c$

$H_1: \mu_i \neq c$



$$\text{Test stat} = \frac{\bar{y}_{i\cdot} - c}{\sqrt{\frac{MSE}{n}}}$$

Then Reject H_0 if $|\text{test stat}| > t_{\alpha/2}$

Consider $\mu_i - \mu_K$

An estimator of $\mu_i - \mu_K$ is $\bar{y}_{i\cdot} - \bar{y}_{K\cdot}$

2

$$V(\bar{y}_{i\cdot} - \bar{y}_{K\cdot}) = V(\bar{y}_{i\cdot}) + V(\bar{y}_{K\cdot}) + 0$$

$$= \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2\sigma^2}{n}$$

So $$\frac{(\bar{y}_{i\cdot} - \bar{y}_{K\cdot}) - (\mu_i - \mu_K)}{\sqrt{2\frac{MSE}{n}}} \sim t_{N-a}$$

Confidence interval for $\mu_i - \mu_K$ is

$$\bar{y}_{i\cdot} - \bar{y}_{K\cdot} \pm t_{\alpha/2}\sqrt{2\frac{MSE}{n}}$$

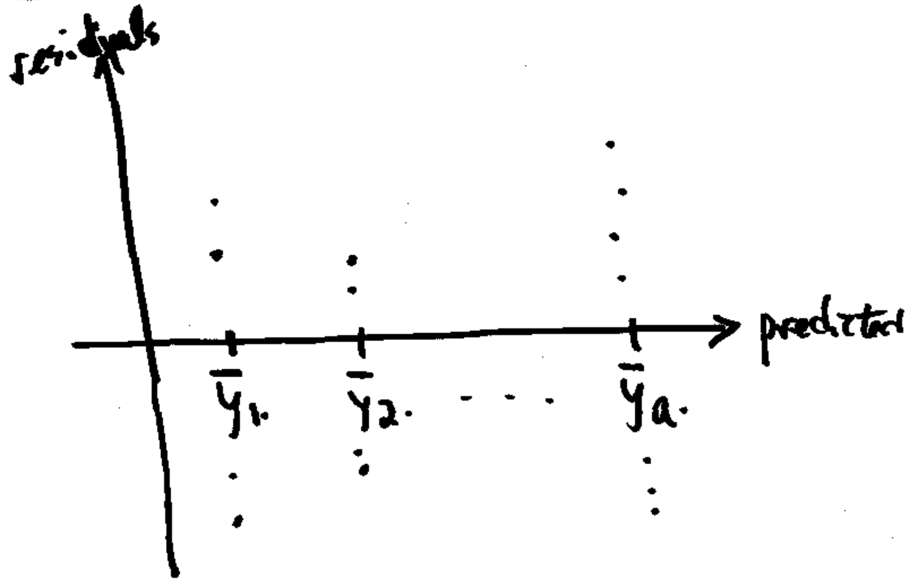
Model verification

① Generate the residuals (standardized or Studentized) + construct a normal probability plot (Q-Q plot or P-P plot)

② Generate a box plot or histogram of the residuals

③ If possible, keep track of the order in which the observations were made, plot the residuals vs. time, + compute the autocorrelation

④ Plot the residuals vs. the predicted values ④



⑤ Levene's or Bartlett's test can be used to check for equality of variances.

(5)

Cautionary note: Levene's & Bartlett's tests are both extremely sensitive to the normality assumption. Also preceding an ANOVA by one of these tests makes the F-test a conditional hypothesis test.

Multiple comparisons

Contrasts: linear combinations of the parameters of the form $\sum_{i=1}^{a} c_i \mu_i$, with $\sum_{i=1}^{a} c_i = 0$

(6)

Examples: $\mu_1 - \mu_2$

$$(\mu_1 - \mu_2) - (\mu_2 - \mu_3) = \mu_1 - 2\mu_2 + \mu_3$$

Let $\Gamma = \sum_{i=1}^{a} c_i \mu_i$ be a contrast

Let $C = \sum_{i=1}^{a} c_i \bar{y}_{i\cdot}$

Then $E[C] = \Gamma$

and $V[C] = \sum_{i=1}^{a} c_i^2 V[\bar{y}_{i\cdot}] + 0$

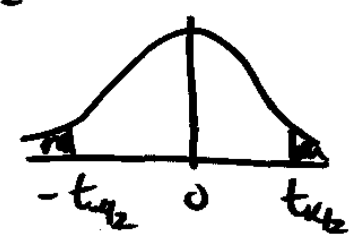
$$= \sum_{i=1}^{a} c_i^2 \frac{\sigma^2}{n} = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$$ 7

Then $$\frac{C - \Gamma}{\sqrt{\frac{MSE}{n} \sum_{i=1}^{a} c_i^2}} \sim t_{N-a}$$

C.I. for Γ : $C \pm t_{\alpha/2} \sqrt{\frac{MSE}{n} \sum_{i=1}^{a} c_i^2}$

H.T. for Γ : $H_0 : \Gamma = 0$

$H_1 : \Gamma \neq 0$



Test stat $= \frac{C - 0}{\sqrt{\frac{MSE}{n} \sum c_i^2}}$

8

Recall:

If $X \sim t_v$

then $X^2 \sim F_{1,v}$

Back to our contrasts,

Under H_0, $$\frac{C - 0}{\sqrt{\frac{MSE}{n} \sum c_i^2}} \sim t_{N-a}$$

So $$\frac{C^2}{\frac{MSE}{n} \sum c_i^2} \sim F_{1,N-a}$$

$$\frac{\frac{C^2}{\frac{1}{n}\sum c_i^2}}{MSE} \sim F_{1, N-a}$$

Let $SS_C = \dfrac{C^2}{\frac{1}{n}\sum c_i^2}$ Under H_0,

$$SS_C \sim \chi^2_1$$

so $MS_C = \dfrac{SS_C}{1}$

<u>Orthogonal</u> <u>contrasts</u>

Let $\vec{c}$ and $\vec{d}$ be the vectors of coefficients

10

for 2 different contrasts

<u>Defn</u>: The contrasts are <u>orthogonal</u> if $\vec{c} \perp \vec{d}$

$$\left(\sum_{i=1}^{a} c_i d_i = 0 \right)$$

<u>Example</u>: We are comparing $\mu_1, \mu_2, \mu_3, \mu_4$

We wish to test:

$\mu_2 = \mu_1$	$\mu_1 - \mu_2 = 0$
$\mu_3 = \dfrac{\mu_1 + \mu_2}{2}$	$\mu_1 + \mu_2 - 2\mu_3 = 0$
$\mu_4 = \dfrac{\mu_1 + \mu_2 + \mu_3}{3}$	$\mu_1 + \mu_2 + \mu_3 - 3\mu_4 = 0$

11

$$\vec{c} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \vec{d} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} \quad \vec{e} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix}$$

Note $\vec{c} \perp \vec{d}$, $\vec{c} \perp \vec{e}$, $\vec{d} \perp \vec{e}$

Also, let $\vec{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Note $\vec{c} \perp \vec{1}$, $\vec{d} \perp \vec{1}$, $\vec{e} \perp \vec{1}$

12

$\therefore$ If there are $\underline{a}$ means being compared, the you can construct at most $a-1$ orthogonal contrasts.

Cochran's Theorem implies that a set of $(a-1)$ orthogonal contrasts will partition SS_{TRT} into $(a-1)$ 1-df sums of squares, leading to $(a-1)$ independent F tests.

(13)

<u>Multiple Comparison methods</u>

- Scheffé's Method for simultaneously comparing <u>all</u> possible contrasts.

Let $\Gamma_K = C_{1K}\mu_1 + \dots + C_{aK}\mu_a$

$K = 1, \dots, \infty$

Γ_K is estimated by

$$C_K = C_{1K}\bar{y}_{1\cdot} + \dots + C_{aK}\bar{y}_{a\cdot}$$

(14)

Scheffé's Theorem says

$$C_K \pm \sqrt{(a-1)F}\sqrt{\frac{MSE}{n}\sum_{i=1}^{a} C_{ik}^2}$$

$F_{\alpha, a-1, N-a}$

provides a joint $(1-\alpha)100\%$ confidence region for <u>all</u> Γ_K, $K = 1, \dots, \infty$

- Fisher's Least Significant Difference Method (LSD)

for pairwise comparisons of μ_i, μ_k

Compute $LSD = t_{\alpha/2} \sqrt{2\frac{MSE}{n}}$

$\uparrow$ $N-a$ df

Then, if $|\bar{y}_{i\cdot} - \bar{y}_{k\cdot}| > LSD$,

declare them significantly different.