STA 303 / 1002

Note Title 1/23/2012

Assignment 1 is ready.

Chunyi T.A. in RW 107 / 109

Mon. Jan. 30 2:00 - 4:00 p.m.

Fri. Feb. 3 2:00 - 4:00 p.m.

General Linear Model for One-way Classification

a.k.a One-way Analysis of Variance

G groups $\quad$ $G-1$ indicator variables in model.

$$I_{g,i} = \begin{cases} 1 & \text{if } i^{th} \text{ observation is in group } g \\ 0 & \text{otherwise} \end{cases}$$

$g = 1, \ldots, G-1$

Model: $Y_i = \beta_0 + \beta_1 I_{1,i} + \cdots + \beta_{G-1} I_{G-1,i} + e_i$

Assumptions: Gauss-Markov conditions:

$E(e_i) = 0$

$Var(e_i) = \sigma^2$ (constant)

$E(e_i e_j) = 0 \quad i \neq j$

e_i's normally distributed

Test that the mean of Y_i is the same for all G group

$H_0: \beta_1 = \beta_2 = \cdots = \beta_{G-1} = 0$

vs H_a: at least one of these is not 0

This is the regression analysis of variance F test

Test statistic: $F_{obs} = \dfrac{SS_{Reg} / G-1}{RSS / \cancel{n-G}\ N-G}$ ← mean square of regression; ← mean square error

Under H_0, F_{obs} is an observation from an F distribution with $G-1$, $\cancel{n-G}$ $N-G$ df.

Correction: "n" should be "N".

In SAS proc glm, this test is given in ANOVA table and under Type III SS.

<u>Spock Example Continued</u>:

Y_i = % of women on i^{th} venire

<u>Now</u> Consider all 7 judges in one model

Evidence of differences among mean of Y_i among 7 judges?

<u>Yes</u> $p < 0.0001$ (test stat: $F_{obs} = 6.72$ use $F_{6,39}$ distribution)

Which judges have different means?

Post-hoc consideration of significant result from analysis of variance F-test

We know that means for all 7 judges are not equal?
Which differ?

One approach: Consider all possible pairs of means
& do 2-sample t-tests
using pooled estimate of error variance
from analysis of variance — calculated
from observations for all 7 judges;
gives better est of variance than just
using observations for the 2 groups

and more df for error $\Rightarrow$ more power

That is, test $H_0: \mu_a = \mu_b$ vs $H_a: \mu_a \neq \mu_b$
for $a, b \in \{1, \ldots, G\}$
for $\binom{G}{2}$ pairs of means

Test stat: $\dfrac{\bar{y}_a - \bar{y}_b}{s\sqrt{\frac{1}{n_a} + \frac{1}{n_b}}}$ where $s = \sqrt{MSE}$

Under H_0, this test stat has a t_{N-G}
if $\mu_a = \mu_b$

or construct $\binom{G}{2}$ CIs for difference in the pairs of means

In <u>SAS</u> <u>means</u> groupvariable / cldiff (CIs)

pdiff (t-tests)

;

$\bar{y}_g$

<u>OR</u> "Least Squares Means"

- means from calculated from fitted model.

eg $\hat{E}(Y_i) = b_0 + b_g$ if i^{th} observation is in g^{th} group

For one way classification, means = least squares means

<u>SAS</u>: lsmeans groupvariable / pdiff;

A problem?

Conducting $\binom{g}{2}$ tests, chosen because of a suggestion from data

(significant F-test from Analysis of variance)

Chance of making at least one Type I error when carrying out many tests is high

Two possible solutions: controlling Type I error rate

① BONFERRONI METHOD: Bonferroni inequality: $P(A \cup B) \le P(A) + P(B)$

Let A_i be event that i^{th} test results in a Type I error

$$P(\cup A_i) \le \sum P(A_i)$$

(correction: Sum was missing)

P(at least Type I error)

In general, k CIs or doing k tests at level α, using α/k, then CI coverage rate or Type I error rate is at most $100(1-\alpha)\%$ / α

That is, chance that at least 1 CI misses what its trying to capture or at least 1 test results in a Type I error is at most α

SAS: bon (means statement) or adjust=bon (lsmeans statement)
option

This multiplies p-values by $\binom{5}{2}$

Bonferroni method is conservative, i.e. overall Type I error rate (chance of making at least one Type I error) is much lower α
(so you're likely to make type II errors.)

② <u>Tukey's Procedure</u>

"Tukey's HSD"   HSD = Honestly significant difference

- usually less conservative than Bonferroni's method, particularly if group sample sizes are (close to) equal

Based on "Studentized Range Distribution" which is based on $\max_{a,b \in \{1,\dots,G\}} \{ \bar{y}_a - \bar{y}_b \}$

Tukey's method simultaneous type I error rate of α

or simultaneous CI coverage rate of $100(1-\alpha)\%$

If sample sizes are unequal, need Kramer adjustment which is conservative.

SAS: tukey option on means statement
adjust=tukey option on lsmeans statement

Conclusion (same for both Bonferroni & Tukey adjustments)

We have evidence that mean % women on venires is different between Spock's judge & all other judges except D ($n_D = 2$) and no evidence of any difference among other judges.

Any <u>problems with model assumptions</u>:

- looked at diagnostics panel
  - one slightly unusual observation but not too bad (large negative residual)
  - variance constant? – look at calculated s.d.
  - normality – OK

- uncorrelated errors – ok if venires chosen independently

Rule of thumb for constant variance assumption:

- calc. s.d. of observations in each group

If $\dfrac{\text{largest s.d.}}{\text{smallest s.d.}} < 2$ ok to proceed as if variances are equal

For Spock's judge example

$\dfrac{\text{largest s.d.}}{\text{smallest s.d}} = \dfrac{11.9}{4.6}$ (ignoring judge D since $n_D = 2$)

so we might have a problem

Consider all inferences as only approximate.

— End of Epoch —

<u>Terminology from Design of</u> Experiments

Factor - categorical predictor variable, often treatments

Factors are composed of <u>levels</u>, e.g. various types of treatments

One-way <u>ANOVA</u> - single factor with 2 or more levels used to explain mean of a continuous response

Two-way ANOVA

- continuous response
- 2 factors