Linear Combination of Means

Design of Experiments - Montgomery
Section 3-5

Linear Combination of Means

$$\begin{aligned} y_{ij} &= \mu + \tau_i + \epsilon_{ij} \\ &= \mu_i + \epsilon_{ij} \end{aligned}$$

- Often hypotheses different than H_0 : all $\tau_i = 0$
- Would like to test $H_0 : L = \sum c_i \mu_i = L_0$
  - Pairwise comparisons
  - Treatments vs control
  - Comparing combinations of trts
  - (Curvi)linear relationship between means and trts
- Hypotheses may be planned or "after the fact"

Can use statistical model to construct t-test (F-test)

$$\widehat{L} = \sum c_i \overline{y}_{i.} \qquad \begin{aligned} \text{Var}(\widehat{L}) &= \text{Var}(\sum c_i \overline{y}_{i.}) \\ &= \sum c_i^2 \text{Var}(\overline{y}_{i.}) \\ &= \text{MS}_\text{E} \sum (c_i^2/n_i) \end{aligned}$$

$$t_o = \frac{\widehat{L} - L_0}{\sqrt{\text{Var}(\widehat{L})}}$$

Under H_0: $t_o \sim t_{N-a}$

Linear Combination of Means

- Why t-test instead of overall F test?
  - T-test specifically addresses hypothesis of interest
  - Overall F-test jointly tests all possible contrasts
  - Why include contrasts of no interest in a test?

    If overall error rate equal to α

    Error rate for single comparison $< \alpha$

    F test will reduce power ("water down" test)

    May find individual test significant while F not
- Problems with linear combinations
  - Multiple tests inflate overall error rate
  - Not all combinations independent
- Will discuss procedures that control error rates

Contrasts

- Linear combination with coeff's that sum to zero
- $\Gamma = \sum c_i \mu_i = 0$ with $\sum c_i = 0$

  To compare Trt1 and Trt2 : $c_i = \{1, -1, 0, \ldots, 0\}$

  To compare Trt1 vs .5(Trt2+Trt3) : $c_i = \{1, -.5, -.5, 0, \ldots, 0\}$
- Will estimate Γ using $C = \sum c_i \overline{y}_{i.}$
- Notice estimate uses trt **means** not totals
- Recall under H_0: $t_0 = C/\sqrt{\text{Var}(C)} \sim t_{N-a}$ (Slide 6-1)
- Also $t^2_{N-a} = F_{1,N-a}$ so could present as F test
- Contrasts often presented in terms of Sum of Squares

  $\text{SS}_C = (\sum c_i \overline{y}_{i.})^2 / \sum (c_i^2/n_i)$

  If divide by MS_E, simply t_0^2

  Can then compare to $F_{1,N-a}$

SAS Procedures (cont.sas)

Example 3-6

```
options ls=80;

title1 'Contrast Comparisons';

data one;
 infile 'c:\saswork\data\tensile.dat';
 input percent strength time;

proc glm data=one;
 class percent;
 model strength=percent;
 contrast 'C1' percent 0 0 0 1 -1;
 contrast 'C2' percent 1 0 1 -1 -1;
 contrast 'C3' percent 1 0 -1 0 0;
 contrast 'C4' percent 1 -4 1 1 1;
-----------------------------------------------------------
Dependent Variable: STRENGTH
                                  Sum of          Mean
Source                  DF       Squares        Square  F Value    Pr > F
Model                    4     475.76000     118.94000    14.76    0.0001
Error                   20     161.20000       8.06000
Corrected Total         24     636.96000

                  R-Square          C.V.      Root MSE      STRENGTH Mean
                  0.746923      18.87642        2.8390             15.040

Source                  DF     Type I SS   Mean Square  F Value    Pr > F
PERCENT                  4     475.76000     118.94000    14.76    0.0001

Contrast                DF   Contrast SS   Mean Square  F Value    Pr > F
C1                       1     291.60000     291.60000    36.18    0.0001
C2                       1      31.25000      31.25000     3.88    0.0630
C3                       1     152.10000     152.10000    18.87    0.0003
C4                       1       0.81000       0.81000     0.10    0.7545
```


Orthogonal Contrasts

- Suppose you have two contrasts $\{c_i\}$ and $\{d_i\}$
- **Orthogonal** if $\sum c_i d_i / n_i = 0$ (using trt means)

- Can divide up SS_{Trt} into $a-1$ orthogonal contrasts
- By Cochran's Thm $\rightarrow$ comparisons independent
- Thus, contrasts of Example 3-6 are independent

- Can also use orthogonal contrasts to study trend
- Only interesting if treatments quantitative (ordered)
- If equally spaced trts and $n_i = n$, c_i in Table X
- Results in breakdown of polynomial regression

$$\mu_i = \beta_0 + \beta_1 i + \ldots + \beta_{a-1} i^{a-1}$$

Determining orthogonal polynomial coefficients using SAS

- Often the levels of the trt are not equally spaced
- Can use Proc IML to determine coeffs

```
proc iml;
levels={1 2 5 10 20};    /* Consider these 5 levels */
print levels;
coef=ORPOL(levels,3);    /* Gives coefs up through cubic */
coef=t(coef);            /* Puts coefs in rows instead of cols */
coef=coef[2:4,];         /* Eliminates the first row of coef matrix*/
print coef;              /* 1st row linear, 2nd quadratic, 3rd cubic */
run;
--------------------------------------------------------------------

                        LEVELS

        1         2         5         10        20


                         COEF

-0.424967 -0.360578 -0.167411 0.1545335  0.798423
0.4348974 0.2072899 -0.325207 -0.711616 0.3946361
-0.433125 0.1365799 0.7252914 -0.510844 0.0820972
```


Testing Multiple Contrasts

- If m orthogonal (independent) contrasts

  P(at least one type I error) $= 1 - (1-\alpha')^m$

  Can control overall error rate using equation

  Example: Consider $m = 5$ independent tests and overall error rate 0.05. For this to hold, each test must use $\alpha' = 1 - (1 - .05)^{1/5} = 0.0102$

- Scheffé's Method

  Set up $1-\alpha$ simultaneous CI for all contrasts

  Protects for unplanned comparisons

  Overall error rate at most $\alpha \leftrightarrow$ low power

  Assumes $SS_C/\sigma^2 \sim \chi^2_{a-1}$ instead of χ^2_1

  Compare $|C|$ to $s_C\sqrt{(a-1)F_{\alpha,a-1,N-a}}$

  Relationship with overall F test

  If P-value of F-test is γ, then at γ level, Scheffé method will only find one contrast significant

$$c_i = n_i(\overline{y}_{i.} - \overline{y}_{..}) = n_i \hat{\tau}_i$$

- Bonferroni's Method

  Recall P($A \cup B$) = P(A) + P(B) - P($A \cap B$)

  P(at least one type I error in m tests) $\leq m\alpha'$

  For each test use $\alpha' = \alpha/m$

  Designed for planned comparisons only

  Looks only at subset of contrasts

  Extremely conservative if m is large

Comparison of Means

- Often only interested in pairwise comparisons
- Can be expressed as contrasts
- If comparing trt j and trt k

$$c_i = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

- Pairwise comparisons a subset of all contrasts
- Want test that considers only subset
- Trade-off between power and prob(Type I error)
- No one "best" test
- Comparison methods vary in protection
  - Experimentwise error (overall Type I)
  - Comparisonwise error (individual Type I)

Pairwise Comparison Methods

- Least Significant Difference (LSD)
  - Use MS_E and df and performs usual t-test

$$\frac{\overline{y}_{i.} - \overline{y}_{j.}}{\sqrt{\text{MS}_\text{E}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \sim t_{N-a}$$

  - Does not control experimentwise error, controls α'
- Protected Least Significant Difference (LSD)
  - Only do comparisons if F-test significant
- Tukey's (Tukey-Kramer) Method
  - Uses studentized range distribution q instead of t
  - Distribution based on $\overline{y}_\text{MAX} - \overline{y}_\text{MIN}$ in a trts
  - Accounts for any possible pair being selected
  - Controls overall experimentwise error rate α
  - Reject if (q available in Table VIII)

$$|\overline{y}_{i.} - \overline{y}_{j.}| > \frac{q_{\alpha,a,N-a}}{\sqrt{2}}\sqrt{\text{MS}_\text{E}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Pairwise Comparison Methods

- Newman-Keuls Test
  - Performs Tukey's test for varying number of trts
  - Start with $p = a$, then $p = a - 1$, etc

$$K_p = q_\alpha(p, f)\sqrt{MS_E/n}$$

  - Stop whenever difference not found significant
  - Controls experimentwise error for all tests with m means
  - When unequal sample size $n = a/\sum 1/n_i$
- Duncan's Multiple Range Test
  - Similar testing approach as Newman-Keuls
  - Based on least significant range (Table VII)
  - Powerful $\leftrightarrow$ does not control overall Type I
- Dunnett's Test
  - Specifically designed for trts vs control situation
  - Distribution based on MAX($\overline{y}_\text{Control} - \overline{y}_\text{Trt}$) for $a - 1$ trts
  - Similar in approach to Tukey's test
  - Controls experimentwise error rate

SAS Procedures

```
options ls=80;

title1 'Means Comparison';

data one;
 infile 'c:\saswork\data\tensile.dat';
 input percent strength time;


proc glm data=one;
 class percent;
 model strength=percent;
 means percent / alpha=.05 lines bon snk tukey;
 means percent / lines duncan lsd scheffe;
 means percent /dunnett;
 means percent / lsd clm;

run;
```



```
            T tests (LSD) for variable: STRENGTH

NOTE: This test controls the type I comparisonwise error rate not
      the experimentwise error rate.

               Alpha= 0.05  df= 20  MSE= 8.06
                 Critical Value of T= 2.09
            Least Significant Difference= 3.7455

Means with the same letter are not significantly different.

        T Grouping              Mean      N  PERCENT

                 A            21.600      5  30

                 B            17.600      5  25
                 B
                 B            15.400      5  20

                 C            10.800      5  35
                 C
                 C             9.800      5  15
```



```
      Bonferroni (Dunn) T tests for variable: STRENGTH

 NOTE: This test controls the type I experimentwise error rate,
      but generally has a higher type II error rate than REGWQ.

               Alpha= 0.05  df= 20  MSE= 8.06
                 Critical Value of T= 3.15
           Minimum Significant Difference= 5.6621

 Means with the same letter are not significantly different.

       Bon Grouping              Mean      N  PERCENT

                  A            21.600      5  30
                  A
          B       A            17.600      5  25
          B
          B       C            15.400      5  20
                  C
                  C            10.800      5  35
                  C
                  C             9.800      5  15
```



```
           Scheffe's test for variable: STRENGTH

NOTE: This test controls the type I experimentwise error rate but
      generally has a higher type II error rate than REGWF for
      all pairwise comparisons

               Alpha= 0.05  df= 20  MSE= 8.06
                Critical Value of F= 2.86608
           Minimum Significant Difference= 6.0796

Means with the same letter are not significantly different.

     Scheffe Grouping              Mean      N  PERCENT

                    A            21.600      5  30
                    A
            B       A            17.600      5  25
            B
            B       C            15.400      5  20
                    C
                    C            10.800      5  35
                    C
                    C             9.800      5  15
```

Tukey's Studentized Range (HSD) Test for variable: STRENGTH

NOTE: This test controls the type I experimentwise error rate, but generally has a higher type II error rate than REGWQ.

Alpha= 0.05 df= 20 MSE= 8.06
Critical Value of Studentized Range= 4.232
Minimum Significant Difference= 5.373

Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	PERCENT
	A	21.600	5	30
	A			
B	A	17.600	5	25
B				
B	C	15.400	5	20
	C			
D	C	10.800	5	35
D				
D		9.800	5	15

Student-Newman-Keuls test for variable: STRENGTH

NOTE: This test controls the type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

Alpha= 0.05 df= 20 MSE= 8.06

Number of Means	2	3	4	5
Critical Range	3.7454541	4.5427098	5.0256318	5.3729606

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	PERCENT
A	21.600	5	30
B	17.600	5	25
B			
B	15.400	5	20
C	10.800	5	35
C			
C	9.800	5	15

Duncan's Multiple Range Test for variable: STRENGTH

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha= 0.05 df= 20 MSE= 8.06

Number of Means	2	3	4	5
Critical Range	3.745	3.931	4.050	4.132

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	PERCENT
A	21.600	5	30
B	17.600	5	25
B			
B	15.400	5	20
C	10.800	5	35
C			
C	9.800	5	15

Dunnett's T tests for variable: STRENGTH

NOTE: This tests controls the type I experimentwise error for comparisons of all treatments against a control.

Alpha= 0.05 Confidence= 0.95 df= 20 MSE= 8.06
Critical Value of Dunnett's T= 2.651
Minimum Significant Difference= 4.7602

Comparisons significant at the 0.05 level are indicated by '***'.

PERCENT Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
30 - 15	7.040	11.800	16.560	***
25 - 15	3.040	7.800	12.560	***
20 - 15	0.840	5.600	10.360	***
35 - 15	-3.760	1.000	5.760	

T Confidence Intervals for variable: STRENGTH

Alpha= 0.05 Confidence= 0.95 df= 20 MSE= 8.06
Critical Value of T= 2.09
Half Width of Confidence Interval= 2.648434

PERCENT	N	Lower Confidence Limit	Mean	Upper Confidence Limit
30	5	18.952	21.600	24.248
25	5	14.952	17.600	20.248
20	5	12.752	15.400	18.048
35	5	8.152	10.800	13.448
15	5	7.152	9.800	12.448