ORIGIN ≡ 0

W. Stein

Simultaneous Comparisons in ANOVA

In simple statistical tests (such as a t-Test), Confidence Intervals may be specified to indicate values of the test statistic in comparison with Critical Values (derived from the inverse cumulative probability t function qt) over which H_0 will *not* be rejected, or equivalently, values of probability greater than a previously specified α.

In using ANOVA, however, an important complication arises. In two population t-Tests, only a single comparison between population means (μ_1 with μ_2) is possible. In ANOVA, however, multiple pairwise or linear contrast comparisons are often of interest. In most cases, these comparisons are made simultaneously, and are therefore dependent upon the same sample data. The existence of multple dependent probabilities implies that the *joint probabity* of a *family of simultaneous comparisons* is greater than probabilities of each comparison separately. If one specifies $\alpha = 0.05$ for one confidence interval (or test) then *familywise* α for all together is always greater (i.e., less significant) than α. Good procedure requires taking familywise probabilities into account especially when presenting results. The example comes from Chapter 16 in Kuter et al. (KNNL) A*pplied Linear Statistical Models* 5th Edition.

Example:

Cell Means ANOVA Model: KNNL p.683

< original variable called "Design" has been dummy coded in this dataset as a set of index variables with: p=4 variables for 4 factor levels.

K := READPRN("c:/2008LinearModelsData/KentonFoodCM.txt")

Variable Assignment:

$Y := K^{\langle 0 \rangle}$

$X1 := K^{\langle 1 \rangle}$ $X2 := K^{\langle 2 \rangle}$ $X3 := K^{\langle 3 \rangle}$ $X4 := K^{\langle 4 \rangle}$

N := length(Y) N **< total number of cases**

X := augment(X1, X2, X3, X4) **< design matrix**

r := cols(X) r = 4

p := r **< p used previously, r = p to conform with KNNL**

i := 0 .. N − 1

ii := 0 .. N − 1

I := identity(N) $J_{i,ii} := 1$ **< Identity & One Matrix for matrix calculations**

$$K = \begin{pmatrix} 11 & 1 & 0 & 0 & 0 \\ 17 & 1 & 0 & 0 & 0 \\ 16 & 1 & 0 & 0 & 0 \\ 14 & 1 & 0 & 0 & 0 \\ 15 & 1 & 0 & 0 & 0 \\ 12 & 0 & 1 & 0 & 0 \\ 10 & 0 & 1 & 0 & 0 \\ 15 & 0 & 1 & 0 & 0 \\ 19 & 0 & 1 & 0 & 0 \\ 11 & 0 & 1 & 0 & 0 \\ 23 & 0 & 0 & 1 & 0 \\ 20 & 0 & 0 & 1 & 0 \\ 18 & 0 & 0 & 1 & 0 \\ 17 & 0 & 0 & 1 & 0 \\ 27 & 0 & 0 & 0 & 1 \\ 33 & 0 & 0 & 0 & 1 \\ 22 & 0 & 0 & 0 & 1 \\ 26 & 0 & 0 & 0 & 1 \\ 28 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$n := \begin{pmatrix} 5 \\ 5 \\ 4 \\ 5 \end{pmatrix}$$ **< Cell numbers counted by hand from K**

Least Squares Estimation of the Regression Parameters:

$$b := \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot Y \qquad b = \begin{pmatrix} 14.6 \\ 13.4 \\ 19.5 \\ 27.2 \end{pmatrix} \qquad \begin{pmatrix} \mu 1 \\ \mu 2 \\ \mu 3 \\ \mu 4 \end{pmatrix}$$ **< estimating μ parameters in the model**

^ vector of regresion coefficients = cell (block) means

Note: In the Cell Means ANOVA model, regression coefficients in the Linear Model give the estimate of means for each cell (block).

Fitted Values & Hat Matrix H:

$Y_h := X \cdot b$ **< fitted values Yh**

$H := X \cdot \left(X^T \cdot X\right)^{-1} \cdot X^T$ **< nXn Hat matrix**

Residuals:

$e := Y - Y_h$ **< residuals**

ANOVA Table:

Sum of Squares:		Degrees of Freedom:		Mean Squares:	
$SSTR := Y^T \cdot \left[H - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTR = (588.2211)$	$df_R := p - 1$	$df_R = 3$	$MSTR := \frac{SSTR}{df_R}$	$MSTR = (196.0737)$
$SSE := Y^T \cdot (I - H) \cdot Y$	$SSE = (158.2)$	$df_E := N - p$	$df_E = 15$	$MSE := \frac{SSE}{df_E}$	$MSE = (10.5467)$
$SSTO := Y^T \cdot \left[I - \left(\frac{1}{N}\right) \cdot J\right] \cdot Y$	$SSTO = (746.4211)$	$df_T := N - 1$	$df_T = 18$	$MSTO := \frac{SSTO}{df_T}$	$MSTO = (41.4678)$

Note: The ANOVA table is calculated in exactly the same way as for regression. SSR now becomes SSTR (for treatment), both with df = 3.

Overall F-Test:

Hypotheses:

H_0: **μ_i the same *for all* i**
$\tau_i = 0$ *for all* i **< Alternate Cell means & Treatment Effects formulations of the same hypothesis**

H_1: ***At least one* μ_i different**
***At least one* τ_i <> 0**

Test Statistic:

$F := \frac{MSTR_0}{MSE_0}$ $F = 18.5911$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$CV := qF[1 - \alpha, (r - 1), (N - r)]$ $CV = 3.2874$

Decision Rule:

IF F > C, THEN REJECT H_0 OTHERWISE ACCEPT H_0

Probability Value:

$P := 1 - pF(F, r - 1, N - r)$ $P = 2.585 \times 10^{-5}$

Tests of Pairs of Means:

Hypotheses:

H_0: $\mu_i = \mu_j$ *for specific* i & j
$\tau_i = \tau_j$ *for specific* i & j

H_1: $\mu_i <> \mu_j$ *for specific* i & j
$\tau_i <> \tau_j$ *for specific* i & j

< Alternate Cell means & Treatment Effects formulations of the same hypothesis

Comparison stated as a Linear Combination of Means:

$$b = \begin{pmatrix} 14.6 \\ 13.4 \\ 19.5 \\ 27.2 \end{pmatrix}$$ < cell means

$$c := \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$ < Contrast Vector = Coefficients of Linear Combination L

$L := \left(b^T \cdot c\right)_0$ $L = 1.2$ < $L = c_1 Ybar_1 + c_2 Ybar_2 + c_3 Ybar_3 + ... + c_k Ybar_k$

Linear Combination with Further Condition as Linear Contrast:

$\sum c = 0$ < Coefficients of the Linear Combination must add to zero to be a Linear Contrast

Single Test of Contrast Vector:

t-Test for Linear Contrast H_0: L = 0 versus H_1: L <> 0:

Cell Means Model:

$Y_{i,j} = \mu_i + \varepsilon_{i,j}$

Factor Effects Model:

$Y_{i,j} = \mu. + \tau_i + \varepsilon_{i,j}$

Equivalent Models, where:
$\mu.$ is the grand mean of all objects
μ_i is mean of each treatment class i (cell or block)
τ_i is the treatment effect = $\mu. - \mu_i$ for each class i.
$\varepsilon_{i,j}$ is the error term specific to each object i,j
i = 1 to r, j = 1 to n_r for each treatment class i

Assumption:

ε_{ij} are a random sample ~ $N(0,\sigma^2)$

For Linear Contrast:

$L = \mu_i c_i$

where:
μ is the vector of Class means
c is some specified contrast vector

Restriction:

$\sum c = 0$

Hypotheses:

H_0: L = 0 < Linear Contrast is zero
H_1: L <> 0

Test Statistic:

$i := 0 .. r - 1 \qquad r = 4$

$$t := \frac{L}{\sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i} \right]}} \qquad t = (0.5842)$$

< Linear Contrast normalized by Standard Error & and Cell sizes

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$C := qt\left(\frac{\alpha}{2}, N - r\right) \qquad C = -2.1314$ **< Note degrees of freedom = (N-r)**

Decision Rule:

IF |t| > |C|, THEN REJECT H_0 OTHERWISE ACCEPT H_0

Probability Value:

$P := \min[\, 2 \cdot pt(t, N - r), 2 \cdot (1 - pt(t, N - r))\,] \qquad P = 0.5677$

Confidence Interval of the Linear Contrast:

$$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i} \right]} \qquad s_{L_0} = 2.0539$$

< Standard Error of L

$CI := \left[\left(L - |C| \cdot s_L\right)_0 \;\; L \;\; \left(L + |C| \cdot s_L\right)_0 \right] \qquad CI = (-3.1779 \;\; 1.2 \;\; 5.5779)$

^ upper limit

confirmed KNNL p. 743 **^ point estimate of L**

^ lower limit

Note: this is a Single Confidence Interval, and other compairsons between menas may be calculated in a similar manner. Once multiple comparisons have been calculated, the set can be considered together with a joint statement of probability. The *familywise probability* of Type I error α for each comparison is potentially much higher, and thus, the confidence interval must be wider. From a mathematical perspective, statisticians routinely caution experimenters about the potential pitfalls of data "snooping" (to borrow a term from KNNL). By this, they mean running a large number of simultaneous tests or confidence intervals, and then proceeding to report significant single findings as if discovered outside the context of the others, or worse, as the result of a strategically-chosen *a priori* experimental design. The problem is that if enough simultaneous tests are run, the laws of probability predict that some tests will end up showing significance merely due to chance. This mathematically-based caution is certainly correct. Given this, are more than one significant planned or unplanned result in ANOVA tests to be considered valid or not? Much depends on what exactly is meant by the foundational concept of α in often widely differing theoretical and experimental contexts. Whereas mathematicians might like to draw a bright line between *a priori* and *post hoc*, in experimental practice rarely is the distinction so clear. All experiments exist within a framework of pre-existing literature and laboratory/field practice for data collection. So *of course* biologists regularly engage in data "snooping" in conceiving of problems, designing studies and analyzing results. They could hardly do otherwise.

In my opinion, the a *priori* vs *post hoc* distinction is of interest from both theoretical and practical standpoints, and to be aware of the issues involved makes it possible to construct stronger scientific arguments. The distinction also points to clear limitations in statistical reasoning in the sciences to the extent that all of it must be acknowledged to be nothing more than an *approximation*. If one's data are overwealmingly clear, then difficulties in the approximation don't really matter. However, if the data are unclear, then how one employs the approximation may influence what one might *say within a test*, but not necessarily what one might *conclude*. The take-home message remains the same. If the data are unclear, then *biological interpretation*, and/or *experimental replication*, must necessarily take precedence over statistical methodology.

Tukey Confidence Interval for Pairwise Contrasts of Means:

$\alpha := 0.10$ **< Set Probability of *Simultaneous* Type I Error**

$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i}\right]}$ $s_L = (2.0539)$ **< sd estimate of L, based on Linear c**

$c = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ $L = 1.2$

$Q := qTukey(1 - \alpha, r, N - r)$ **< 'q' function for Studentized Range Distribution - Not present in MathCad**

$Q := 3.5398914$ **< So I set explicitly it here by consulting KNNL Table B.9 or qtukey{stats} in R**

$T := \frac{1}{\sqrt{2}} \cdot Q$ $T = 2.5031$ **< Tukey multiplier**

$CI_T := \left[(L - T \cdot s_L)_0 \;\; L \;\; (L + T \cdot s_L)_0\right]$ $CI_T = (-3.9412 \;\; 1.2 \;\; 6.3412)$

^ upper limit

confirmed KNNL p. 751 **^ point estimate of L = μ1-μ2**

^ lower limit

Note: When used with unequal sample sizes, the Tukey Procedure is also called the Tukey-Kramer Procedure. The Tukey procedure is only to be used with pairwise comparisons of means. The confidence intervals (or equivalent t/F tests) warrent that all pairwise comparisons of means have family (simultaneous) confidence less than specified α (thus conservative). Data "snooping" is allowed so long as comparisons are confined to the family of pair-wise comparison of means. Otherwise, use Scheffe or Bonferroni methods.

Scheffe Multiple Comparison Confidence Interval for all Linear Contrasts:

$\alpha := 0.10$ **< Set Probability of *Simultaneous* Type I Error**

$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i}\right]}$ $s_L = (2.0539)$ **< sd estimate of L, based on Linear contrast c**

$c = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ $L = 1.2$

$S := \sqrt{(r - 1) \cdot qF(1 - \alpha, r - 1, N - r)}$ $S = 2.733$ **< Scheffe multiplier**

$CI_S := \left[(L - S \cdot s_L)_0 \;\; L \;\; (L + S \cdot s_L)_0\right]$ $CI_S = (-4.4134 \;\; 1.2 \;\; 6.8134)$

To Verify a different Linear Contrast in KNNL:

$b = \begin{pmatrix} 14.6 \\ 13.4 \\ 19.5 \\ 27.2 \end{pmatrix}$ **< cell means** $c := \begin{pmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{pmatrix}$ **< Contrast Vector = Coefficients of Linear Combination L**

$L_1 := \left(b^T \cdot c\right)_0$ $L_1 = -9.35$ **< $L_1 = c_1Ybar_1 + c_2Ybar_2 + c_3Ybar_3 + ... + c_kYbar_k$**

$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i}\right]}$ $s_L = (1.4971)$ **< sd estimate of L, based on Linear conrast c** $c = \begin{pmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{pmatrix}$ $L = 1.2$

$CI_S := \left[\left(L_1 - S \cdot s_L\right)_0 \quad L_1 \quad \left(L_1 + S \cdot s_L\right)_0\right]$ $CI_S = (-13.4415 \quad -9.35 \quad -5.2585)$

^ upper limit

confirmed KNNL p. 755 **^ point estimate of L = 0.5(μ1+μ2) - 0.5(μ3+μ4)**

^ lower limit

Note: Scheffe Confidence Intervals (and equivalent t/F tests) provide family (simultaneous) confidence intervals for the entire set of Linear Contrasts - an infinite set. Thus, the test is very conservative and suitable for all forms of data "snooping". However, one might be able to more accurately define family confidence intervals using either Tukey or Bonferroni methods.

Bonferroni Multiple Comparison Confidence Interval for a Set of Linear Contrasts:

$\alpha := 0.10$ **< Set Probability of *Simultaneous* Type I Error**

$b = \begin{pmatrix} 14.6 \\ 13.4 \\ 19.5 \\ 27.2 \end{pmatrix}$ **< cell means** $c := \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ **< Contrast Vector = Coefficients of Linear Combination L**

$L_1 := \left(b^T \cdot c\right)_0$ $L_1 = 1.2$ **< $L_1 = c_1Ybar_1 + c_2Ybar_2 + c_3Ybar_3 + ... + c_kYbar_k$**

$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i}\right]}$ $s_L = (2.0539)$ **< sd estimate of L, based on Linear contrast c** $c = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ $L = 1.2$

$g := 6$ **< number of Contrasts considered = all pairwise combinations of 4 means**

$B := qt\left(1 - \frac{\alpha}{2 \cdot g}, N - r\right)$ $B = 2.6937$ **Bonferroni multiplier**

$CI_B := \left[\left(L - |B| \cdot s_L\right)_0 \quad L \quad \left(L + |B| \cdot s_L\right)_0\right]$ $CI_B = (-4.3328 \quad 1.2 \quad 6.7328)$

To Verify a different Linear Contrast in KNNL:

$\alpha := 0.025$ **< Set as given in KNNL p. 756** $1 - \alpha = 0.975$

$b = \begin{pmatrix} 14.6 \\ 13.4 \\ 19.5 \\ 27.2 \end{pmatrix}$ **< cell means** $c := \begin{pmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{pmatrix}$ **< Contrast Vector = Coefficients of Linear Combination L**

$L_2 := \left(b^T \cdot c\right)_0$ $L_2 = -9.35$ **< $L_1 = c_1 Ybar_1 + c_2 Ybar_2 + c_3 Ybar_3 + ... + c_k Ybar_k$**

$s_L := \sqrt{MSE \cdot \left[\sum_i \frac{(c_i)^2}{n_i}\right]}$ $s_L = (1.4971)$ **< sd estimate of L, based on Linear conrast c** $c = \begin{pmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{pmatrix}$ $L = 1.2$

$g := 2$ **< number of Contrasts considered = all pairwise combinations of 4 means**

$B := qt\left(1 - \frac{\alpha}{2 \cdot g}, N - r\right)$ $B = 2.8366$ **Bonferroni multiplier**

$CI_B := \left[\left(L_2 - |B| \cdot s_L\right)_0 \; L_2 \; \left(L_2 + |B| \cdot s_L\right)_0\right]$ $CI_B = (-13.5966 \quad -9.35 \quad -5.1034)$

^ upper limit

confirmed KNNL p. 757 **^ point estimate of**

^ lower limit **$L = 0.5(\mu 1+\mu 2) - 0.5(\mu 3+\mu 4)$**

Note: Bonferroni family (simultaneous) confidence intervals are defined by g = an explicit number of contrasts set *a priori* (in advance) by the researcher. Thus data "snooping" beyond this set is not permitted.

Prototype in R:

```
TREATMENTS MODEL IN R
#READ STRUCTURED DATA TABLE WITH NUMERIC CODED FACTOR
K=read.table("c:/2008LinearModelsData/KentonFoodR.txt")
K
attach(K)
Y=Sales
X=factor(Design) # factor() IN DEFAULT SETTING
FM=lm(Y~X)
model.matrix(FM)
summary(FM)
anova(FM)
MSE=summary(FM)$sigma^2
MSE
```

```
> anova(FM)
Analysis of Variance Table
Response: Y
          Df Sum Sq Mean Sq F value    Pr(>F)
X          3 588.22  196.07  18.591 2.585e-05 ***
Residuals 15 158.20   10.55
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
'.' 0.1 ' ' 1

> MSE
[1] 10.54667
```

```
aov(FM)
#TUKEYHSD MULTIPLE COMPARISONS
#TUKEYHSD MULTIPLE COMPARISONS
TukeyHSD(aov(FM),conf.level=0.90)
```

```
> TukeyHSD(aov(FM),conf.level=0.90)
  Tukey multiple comparisons of means
    90% family-wise confidence level
Fit: aov(formula = FM)
$X
    diff        lwr       upr     p adj
2-1 -1.2 -6.3411769  3.941177 0.9352978
3-1  4.9 -0.5530416 10.353042 0.1548895
4-1 12.6  7.4588231 17.741177 0.0001013
3-2  6.1  0.6469584 11.553042 0.0582866
4-2 13.8  8.6588231 18.941177 0.0000368
4-3  7.7  2.2469584 13.153042 0.0142180
```

Same, however note difference in sign >

```
#STUDENTIZED RANGE DISTRIBUTION FUNCTIONS
?ptukey
?qtukey
alpha=0.10
N=length(Y) #TOTAL NUMBER OF CASES
levels(X)
r=4      #NUMBER OF LEVELS IN FACTOR X
Q=qtukey(1-alpha,r,N-r)
Q
```

```
> Q=qtukey(1-alpha,r,N-r)
> Q
[1] 3.539891
```

value used in MathCad calculation above ^

```
?pairwise.t.test()

#PAIRWISE COMPARISONS USING BONFERRONI & HOLM
pairwise.t.test(Y,X,"none")  #SINGLE CI'S
pairwise.t.test(Y,X,"bonferroni") #BONFERRONI
pairwise.t.test(Y,X,"holm")  # A USEFUL BONFERRONI VARIANT
```

pairwise single P >

```
> pairwise.t.test(Y,X,"none")  #SINGLE P
        Pairwise comparisons using t tests with pooled SD
data:  Y and X
  1       2       3
2 0.568   -       -
3 0.040   0.013   -
4 1.9e-05 6.9e-06 0.003
P value adjustment method: none
```

pairwise Bonferroni P >

```
> pairwise.t.test(Y,X,"bonferroni") #BONFERRONI P
        Pairwise comparisons using t tests with pooled SD
data:  Y and X
  1       2       3
2 1.00000 -       -
3 0.23969 0.08075 -
4 0.00011 4.1e-05 0.01802
P value adjustment method: bonferroni
```

pairwise Holm P >

```
> pairwise.t.test(Y,X,"holm")  # A USEFUL BONFERRONI VARIANT
        Pairwise comparisons using t tests with pooled SD
data:  Y and X
  1       2       3
2 0.568   -       -
3 0.080   0.040   -
4 9.6e-05 4.1e-05 0.012
P value adjustment method: holm
```

```
#MAKING CONTRASTS IN R
require(gmodels)  #MUST LOAD {gmodels} PACKAGE FROM CRAN

#CONTRAST FOR KNNL P. 743
fit.contrast(FM,X,c(0.5,0.5,-0.5,-0.5),conf.int=0.95)
```

```
> fit.contrast(FM,X,c(0.5,0.5,-0.5,-0.5),conf.int=0.95)
                           Estimate Std. Error   t value     Pr(>|t|)  lower CI  upper CI
X c=( 0.5 0.5 -0.5 -0.5 )    -9.35   1.497053 -6.245605 1.567514e-05 -12.54089 -6.159108
```

```
#SINGLE CALCULATION OF SCHEFFE

alpha=.10  #SET FAMILYWIDE CONFIDENCE ALPHA
N=length(Y) #TOTAL NUMBER OF CASES
levels(X)
r=4       #NUMBER OF LEVELS IN FACTOR X
S=sqrt((r-1)*qf(1-alpha,r-1,N-r)) #SCHEFFE MULTIPLIER
S
#CALCULATING NECESSARY VALUES FROM CONTRAST
LFIT=fit.contrast(FM,X,c(0.5,0.5,-0.5,-0.5),conf.int=0.90)
LFIT
L=LFIT[1] # CALCULATION OF LINEAR CONTRAST L
L
sL=LFIT[2] # CALCULATION OF sL
sL
CI=LFIT[5:6]  # SINGLE CONFIDENCE INTERVAL
CI
CIScheffelower=L-S*sL
CIScheffeupper=L+S*sL
CIS=cbind(CIScheffelower,L,CIScheffeupper)
CIS  #SCHEFFE FAMILY CONFIDENCE INTERVAL
```

```
> CIS  #SCHEFFE FAMILY CONFIDENCE INTERVAL
     CIScheffelower     L CIScheffeupper
[1,]      -13.44147 -9.35      -5.258535
```

Same result for KNNL Linear contrast p. 755, and above ^

```
#CONVERTING P VALUES IN pairwise.t.test() INTO CI'S
options(digits=10)
alpha=0.10

#SINGLE CI'S
C=abs(qt(alpha/2,N-r)) #SINGLE MULTIPLIER
C
summary(FM) #CALCULATE BY HAND: coef +/- C*Std.Error
```

$C := 1.753050$

$L := -1.2$ **< first contrast**

$sL_1 := 2.053939$ **< for first contrast**

$CI_{lwr} := L - C \cdot sL_1$

$CI_{upr} := L + C \cdot sL_1$

$CI := (CI_{lwr} \quad CI_{upr})$

$CI = (-4.8007 \quad 2.4007)$

```
> C
[1] 1.753050

> summary(FM)
Call:
lm(formula = Y ~ X)
Residuals:
   Min     1Q Median     3Q    Max
 -5.20  -1.95  -0.20   1.50   5.80
Coefficients:
             Estimate Std. Error  t value   Pr(>|t|)
(Intercept) 14.600000   1.452354 10.05264 4.6632e-08 ***
X2          -1.200000   2.053939 -0.58424   0.567740
X3           4.900000   2.178532  2.24922   0.039948 *
X4          12.600000   2.053939  6.13455 1.9101e-05 ***
```

```
CS12=fit.contrast(FM,X,c(1,-1,0,0),conf.int=0.90)
CS13=fit.contrast(FM,X,c(1,0,-1,0),conf.int=0.90)
CS14=fit.contrast(FM,X,c(1,0,0,-1),conf.int=0.90)
CS23=fit.contrast(FM,X,c(0,1,-1,0),conf.int=0.90)
CS24=fit.contrast(FM,X,c(0,1,0,-1),conf.int=0.90)
CS34=fit.contrast(FM,X,c(0,0,1,-1),conf.int=0.90)
CS12
CS13
CS14
CS23
CS24
CS34
```

```
> CS12
                Estimate  Std. Error     t value     Pr(>|t|)     lower CI    upper CI
X c=( 1 -1 0 0 )      1.2 2.053939305 0.584243165 0.5677402031 -2.400659028 4.800659028
> CS13
                Estimate  Std. Error      t value      Pr(>|t|)     lower CI     upper CI
X c=( 1 0 -1 0 )     -4.9 2.178531616 -2.249221432 0.03994770436 -8.719075624 -1.080924376
> CS14
                Estimate  Std. Error      t value        Pr(>|t|)     lower CI     upper CI
X c=( 1 0 0 -1 )    -12.6 2.053939305 -6.134553233 1.910148869e-05 -16.20065903 -8.999340972
> CS23
                Estimate  Std. Error      t value      Pr(>|t|)     lower CI     upper CI
X c=( 0 1 -1 0 )     -6.1 2.178531616 -2.800051170 0.01345833690 -9.919075624 -2.280924376
> CS24
                Estimate  Std. Error      t value        Pr(>|t|)     lower CI     upper CI
X c=( 0 1 0 -1 )    -13.8 2.053939305 -6.718796398 6.880921391e-06 -17.40065903 -10.19934097
> CS34
                Estimate  Std. Error      t value       Pr(>|t|)     lower CI     upper CI
X c=( 0 0 1 -1 )     -7.7 2.178531616 -3.534490822 0.003003287344 -11.51907562 -3.880924376
```

CS12 CI is the same as above ^

```
#BONFERRONI CI'S
alpha=0.10
g=6  #NUMBER OF PAIRWISE CONTRASTS
B=qt(1-(alpha/(2*g)),N-r) #BONFERRONI MULTIPLIER
B

lwr=CS12[1]-CS12[2]*B
upr=CS12[1]+CS12[2]*B
CB12=cbind(lwr,upr)
```

```
> B
[1] 2.693739319
```

```
lwr=CS13[1]-CS13[2]*B
upr=CS13[1]+CS13[2]*B
CB13=cbind(lwr,upr)
lwr=CS14[1]-CS14[2]*B
upr=CS14[1]+CS14[2]*B
CB14=cbind(lwr,upr)
lwr=CS23[1]-CS23[2]*B
upr=CS23[1]+CS23[2]*B
CB23=cbind(lwr,upr)
lwr=CS24[1]-CS24[2]*B
upr=CS24[1]+CS24[2]*B
CB24=cbind(lwr,upr)
lwr=CS34[1]-CS34[2]*B
upr=CS34[1]+CS34[2]*B
CB34=cbind(lwr,upr)

CB12
CB13
CB14
CB23
CB24
CB34
```

```
> CB12
              lwr         upr
[1,] -4.332777064 6.732777064
> CB13
              lwr         upr
[1,] -10.76839627 0.968396271
> CB14
              lwr          upr
[1,] -18.13277706 -7.067222936
> CB23
              lwr           upr
[1,] -11.96839627 -0.2316037290
> CB24
              lwr          upr
[1,] -19.33277706 -8.267222936
> CB34
              lwr          upr
[1,] -13.56839627 -1.831603729
```

< same as above

```
#SCHEFFE CI'S
alpha=0.10
S=sqrt((r-1)*qf(1-alpha,r-1,N-r)) #SCHEFFE MULTIPLIER
S

lwr=CS12[1]-CS12[2]*S
upr=CS12[1]+CS12[2]*S
CSF12=cbind(lwr,upr)
lwr=CS13[1]-CS13[2]*S
upr=CS13[1]+CS13[2]*S
CSF13=cbind(lwr,upr)
lwr=CS14[1]-CS14[2]*S
upr=CS14[1]+CS14[2]*S
CSF14=cbind(lwr,upr)
lwr=CS23[1]-CS23[2]*S
upr=CS23[1]+CS23[2]*S
CSF23=cbind(lwr,upr)
lwr=CS24[1]-CS24[2]*S
upr=CS24[1]+CS24[2]*S
CSF24=cbind(lwr,upr)
lwr=CS34[1]-CS34[2]*S
upr=CS34[1]+CS34[2]*S
CSF34=cbind(lwr,upr)

CSF12
CSF13
CSF14
CSF23
CSF24
CSF34
```

```
> S
[1] 2.733013575
```

```
> CSF12
              lwr         upr
[1,] -4.413444002 6.813444002
> CSF13
              lwr         upr
[1,] -10.85395648 1.053956479
> CSF14
             lwr          upr
[1,] -18.213444 -6.986555998
> CSF23
              lwr           upr
[1,] -12.05395648 -0.1460435209
> CSF24
             lwr          upr
[1,] -19.413444 -8.186555998
> CSF34
              lwr          upr
[1,] -13.65395648 -1.746043521
```