Chapter 3 - Kuehl

3.1 TREATMENT COMPARISONS ANSWER RESEARCH QUESTIONS

- Often questions arise Regarding specific treatment means. These may be simply comparing pairs of treatment means, or more complex comparisons among groups of treatment means.

3.2 Linear Functions of observations & Contrasts

Linear Functions of observations (Appendix 3A)

Let $Y_1, \ldots, Y_n$ be a sample of random variables

$c = \sum_{i=1}^{n} k_i Y_i = k_1 Y_1 + \cdots + k_n Y_n$ is a linear function of the random variables.

$$E[Y_i] = \mu_i \qquad V[Y_i] = \sigma_i^2 \qquad Cov(Y_i, Y_j) = \sigma_{ij}$$

$$\mu_c = E(c) = E\left[\sum k_i Y_i\right] = \sum_{i=1}^{n} E[k_i Y_i] = \sum_{i=1}^{n} k_i E(Y_i) = \sum_{i=1}^{n} k_i \mu_i$$

$$\sigma_c^2 = V(c) = V\left[\sum_i k_i Y_i\right] = \sum_{i=1}^{n} k_i^2 V(Y_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} k_i k_j Cov(Y_i, Y_j)$$

$$= \sum_{i=1}^{n} k_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} k_i k_j \sigma_{ij}$$

Special Cases

Sample Mean $\bar{Y} = C = \sum_{i=1}^{n} k_i Y_i = \sum_{i=1}^{n} \frac{1}{n} Y_i$

$$\mu_C = \frac{1}{n} \sum_{i=1}^{n} \mu_i$$

$$\sigma_C^2 = \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 + 2\left(\frac{1}{n^2}\right) \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij}$$

If this represents a sample of independent observations from a Normal distribution with mean μ, variance σ^2:

$$\mu_i = \mu \ \forall\ i, \quad \sigma_i^2 = \sigma^2 \ \forall\ i \quad \sigma_{ij} = 0 \ \forall\ i \neq j$$

AND SAMPLING DIST^{N} OF $\bar{Y}$ is normal:

$$\mu_C = \frac{1}{n} \sum_{i=1}^{n} \mu = \frac{1}{n}(n\mu) = \mu$$

$$\sigma_C^2 = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2}\left[n\sigma^2\right] = \frac{\sigma^2}{n}$$

Linear Functions of Sample Means (Independent Samples)

$\bar{Y}_i \equiv$ Sample mean from population i, based on sample of r_i

$$C = \sum_{i=1}^{t} k_i \bar{Y}_i$$

$$\mu_C = \sum_{i=1}^{t} k_i E\left[\bar{Y}_i\right] = \sum_{i=1}^{t} k_i \mu_i$$

$$\sigma_C^2 = \sum_{i=1}^{t} k_i^2 V(\bar{Y}_i) = \sum_{i=1}^{t} k_i^2 \frac{\sigma_i^2}{r_i}$$

Note: $Cov(\bar{Y}_i, \bar{Y}_j) = 0$ when samples are independent, as they are in the Completely Randomized Design

Contrasts of Means

CONTRASTS OF MEANS ARE LINEAR FUNCTIONS OF MEANS SUCH THAT THE COEFFICIENTS SUM TO 0.

Notation: UPPER CASE LETTERS ⟹ CONTRAST OF POPULATION MEANS.

LOWER CASE LETTERS ⟹ CONTRAST OF SAMPLE MEANS.

$$C = \sum_{i=1}^{t} k_i \mu_i \quad \ni \quad \sum_{i=1}^{t} k_i = 0$$

$$c = \sum_{i=1}^{t} k_i \bar{y}_{i.} \quad \ni \quad \sum_{i=1}^{t} k_i = 0$$

NOTE: $C = \sum_{i=1}^{t} k_i \mu_i$ is an unknown parameter that we wish to make inferences about

ESTIMATOR: $c = \sum_{i=1}^{t} k_i \bar{y}_{i.}$ $\qquad \bar{y}_{i.} \sim N(\mu_i, \frac{\sigma^2}{r_i})$ under current

$$E[c] = E\left[\sum_{i=1}^{t} k_i \bar{y}_{i.}\right] = \sum_{i=1}^{t} k_i E(\bar{y}_{i.}) = \sum_{i=1}^{t} k_i \mu_i$$

$$V[c] = V\left[\sum_{i=1}^{t} k_i \bar{y}_{i.}\right] = \sum_{i=1}^{t} k_i^2 \cdot V[\bar{y}_{i.}] = \sum_{i=1}^{t} \frac{\sigma^2 k_i^2}{r_i}$$

Balanced Data $(r_1 = r_2 = \cdots = r_t = r)$

$$\Rightarrow V(c) = \frac{\sigma^2}{r} \sum_{i=1}^{t} k_i^2$$

Estimated Variance: $S_c^2 = \frac{S^2}{r} \sum_{i=1}^{t} K_i^2$

ESTIMATED STANDARD ERROR: $S_c = \sqrt{S_c^2}$

Interval Estimates for Contrasts

$$C \pm t_{\frac{\alpha}{2}, N-t} \cdot S_c$$

Sum of SQUARES PARTITION FOR CONTRAST

$$SSC = \frac{\left(\sum_{i=1}^{t} K_i \bar{Y}_{i.} \right)^2}{\sum_{i=1}^{t} (K_i^2 / r_i)}$$

if Design is balanced: $SSC = r \frac{\left(\sum_{i=1}^{t} K_i \bar{Y}_{i.} \right)^2}{\sum_{i=1}^{t} K_i^2}$

Hypothesis Tests Regarding Contrasts

$H_0: C = \sum_{i=1}^{t} K_i \mu_i = 0$

$H_A: C \neq 0$

i) F-Test

T.S. $F_0 = \frac{MSC}{MSE} = \frac{SSC/1}{SSE/(N-t)}$

RR: $F_0 \geq F_{\alpha, 1, N-t}$

ii) t-test

T.S. $t_0 = \frac{c}{s_c}$

RR: $|t_0| \geq t_{\frac{\alpha}{2}, N-t}$

Connection Between F-test & t-test

$$F_0 = \frac{MSC}{MSE} = \frac{SSC}{s^2} = r\frac{(\sum k_i \bar{y}_{i.})^2}{s^2 \sum k_i^2}$$

$$t_0 = \frac{c}{s_c} = \frac{\sum k_i \bar{y}_{i.}}{s\sqrt{\sum \frac{k_i^2}{r}}} = \sqrt{r}\,\frac{\sum k_i \bar{y}_{i.}}{s\sqrt{\sum k_i^2}}$$

$$\Rightarrow t_0^2 = F_0$$

Also: $\left[t_{\frac{\alpha}{2}, N-t}\right]^2 = F_{\alpha, N-t}$

Orthogonal Contrasts

Two contrasts are orthogonal if the sum of the products of their coefficients is 0 (when weighted by reciprocal of sample sizes.)

$c = \sum_{i=1}^{t} k_i \bar{y}_{i.}$ $\qquad d = \sum_{i=1}^{t} d_i \bar{y}_{i.}$

c & d are orthogonal if $\sum_{i=1}^{t} \frac{k_i d_i}{r_i} = 0$

Note: if all r_i are $=$, then c, d orthogonal if $\sum_{i=1}^{t} k_i d_i = 0$

<u>NOTE:</u> If we have $t-1$ pairwise orthogonal contrasts, their sums of squares will add up to the treatment sum of squares. That is, we can <u>partition</u> the treatment sum of squares into $t-1$ pairwise orthogonal contrasts.

NEJM - July 11, 2002 - A CONTROLLED TRIAL OF ARTHROSCOPIC SURGERY FOR OSTEOARTHRITIS OF THE KNEE 347: 81-88

<u>EXAMPLE</u> - PLACEBO CONTROLLED TRIAL FOR ARTHROSCOPIC KNEE SURGERY (ASSUMING $r_1 = r_2 = r_3 = r = 54$)

			Actual r
TRT 1 - PLACEBO	$\bar{Y}_{1.} = 53.6$	$S_1 = 22.1$	$r_1 = 54$
TRT 2 - LAVAGE (ACTIVE)	$\bar{Y}_{2.} = 57.8$	$S_2 = 23.5$	$r_2 = 57$
TRT 3 - Débridement (Active)	$\bar{Y}_{3.} = 53.3$	$S_3 = 25.4$	$r_3 = 51$

CASE 1 - TREATING DATA AS BALANCED ($r = 54$)

TRT	r_i	$\bar{Y}_{i.}$	S_i	$r_i(\bar{Y}_{i.} - \bar{Y}_{..})^2$	$(r_i - 1) S_i^2$
1	54	53.6	22.1	91.26	25885.73
2	54	57.8	23.5	454.14	29269.25
3	54	53.3	25.4	138.24	34193.48
		$\bar{Y}_{..} = 54.9$		683.64	89348.46
				$= SST$ TREATMENT SS	$= SSE$ ERROR SS

$$MST = \frac{SST}{t-1} = \frac{683.64}{3-1} = 341.82$$

$$MSE = \frac{SSE}{N-t} = \frac{89348.46}{162-3} = 561.94$$

$$F_0 = \frac{341.82}{561.94} = 0.61 \qquad F_{.05,2,159} \approx 3.07 \text{ (being conservative)}$$

ARTHROSCOPIC SURGERY EXAMPLE CONTINUED

Consider 2 CONTRASTS:

Contrast 1: PLACEBO vs 2 ACTIVES

$$C = 2\mu_1 - \mu_2 - \mu_3$$

(This is equivalent to comparing μ_1 w/ mean of μ_2 & μ_3, but math is easier when working w/ integers)

$K_1 = 2 \quad K_2 = -1 \quad K_3 = -1 \implies \sum K_i = 0$

$$c = 2\bar{y}_{1.} - \bar{y}_{2.} - \bar{y}_{3.} = 2(53.6) - 57.8 - 53.3 = -3.90$$

$$s_c^2 = s^2 \sum_{i=1}^{3} \frac{k_i^2}{r_i} = 561.94\left[\frac{4+1+1}{54}\right] = \frac{561.94}{9} = 62.44$$

$$s_c = \sqrt{62.44} = 7.90$$

95% CI for $C = 2\mu_1 - \mu_2 - \mu_3$

$$c \pm t_{.025,159} \cdot s_c \doteq c \pm z_{.025} \cdot s_c$$

$$\equiv -3.90 \pm 1.96(7.90) \equiv -3.90 \pm 15.49 \equiv \underbrace{(-19.39, 11.59)}_{\text{Contains } 0}$$

$$SSC = r\frac{\left(\sum k_i \bar{y}_{i.}\right)^2}{\sum k_i^2} = 54\frac{(-3.90)^2}{[4+1+1]}$$

$$= 9(-3.90)^2 = 136.89$$

Do F-test & confirm equivalence to t-test.

ARTHROSCOPIC SURGERY EXAMPLE CONTINUED

Contrast 2 – LAVAGE vs. Débridement (comparison of Actives)

$$D = \mu_2 - \mu_3 \qquad d_1 = 0 \quad d_2 = 1 \quad d_3 = -1 \quad \Sigma d_i = 0$$

$$d = \bar{y}_{2.} - \bar{y}_{3.} = 57.8 - 53.3 = 4.5$$

$$s_d^2 = s^2 \sum_{i=1}^{3} \frac{k_i^2}{r_i} = 561.94\left[\frac{0+1+1}{54}\right] = \frac{561.94}{27} = 20.81$$

$$s_d = 4.56$$

95% CI for $\mu_2 - \mu_3$

$$d \pm z_{.025} \cdot s_d \equiv 4.5 \pm 1.96(4.56) \equiv 4.5 \pm 8.94$$

$$= (-4.44, 13.44)$$

Contains 0.

$$SSD = r\frac{(\Sigma d_i \bar{y}_{i.})^2}{\Sigma d_i^2} = 54\frac{(4.5)^2}{(0+1+1)} = 27(4.5)^2 = \text{[redacted]}\ 546.75$$

Conduct F-test and confirm equivalence to [redacted] t-test.

Note: $\sum_{i=1}^{t} k_i d_i = 2(0) + (-1)(1) + (-1)(-1) = 0 - 1 + 1 = 0$

$\Rightarrow$ orthogonal under equal reps "assumption"

$$SSC + SSD = 136.89 + 546.75 = 683.64 = SST$$

(SST has 2 df, we have 2 orthogonal contrasts).

CASE 2 - ACTUAL SAMPLE SIZES (UNBALANCED)

i	K_i	d_i	r_i	$K_i d_i / r_i$
1	2	0	54	$0/54$
2	-1	1	57	$-1/57$
3	-1	-1	51	$1/51$

$$\sum \frac{K_i d_i}{r_i} = \frac{0(57)(51) + (-1)(54)(51) + (1)(54)(57)}{54(57)(51)} =$$

$$= \frac{0 - 2754 + 3078}{54(57)(51)} = \frac{324}{156978} \neq 0$$

To get an orthogonal contrast, start with the second (d) contrast and leave it as

$d = \bar{Y}_{2.} - \bar{Y}_{3.}$ $\quad (d_1 = 0,\ d_2 = 1,\ d_3 = -1)$

- Now select the coefficients of the first contrast that have the same weights as their sample sizes (e.g. $K_2 = 57$, $K_3 = 51$). Since these are divisible by 3, we can use $K_2 = 19$ and $K_3 = 17$
- Finally choose K_1 so that the coefficients for contrast 1 sum to 0 (definition of contrast). This implies $K_1 = -(57+51) = -108$. We could also have had $K_1 = 108$, $K_2 = -57$, $K_3 = -51$
- Checking for orthogonality ($K_1 = 108$, $K_2 = -57$, $K_3 = -51$)

$$\sum_{i=1}^{3} \frac{K_i d_i}{r_i} = \frac{108(0)}{54} + \frac{(-57)(1)}{57} + \frac{(-51)(-1)}{51}$$

$$= 0 - 1 + 1 = 0 \checkmark$$

ANOVA for Unbalanced Data (Actual Sample Sizes)

i	r_i	$\bar{Y}_{i.}$	S_i	$r_i(\bar{Y}_{i.}-\bar{Y}_{..})^2$	$(r_i-1)S_i^2$
1	54	53.6	22.1	103.335	25885.73
2	57	57.8	23.5	452.216	30926.00
3	51	53.3	25.4	144.514	32258.00
		$\bar{Y}_{..} = 54.98\bar{3}$		SST = 700.065	SSE = 89069.73

$$\bar{Y}_{..} = \frac{54(53.6)+57(57.8)+51(53.3)}{54+57+51} = \frac{8907.3}{162} = \text{[redacted]}\ 54.98\bar{3}$$

- NOTE slight changes in SST, SSE from having treated data as balanced. would have been much larger had the degree of unbalance been higher.

- $C_\bullet = \sum k_i \bar{Y}_{i.} = 108\bar{Y}_{1.} - 57\bar{Y}_{2.} - 51\bar{Y}_{3.}$

$$= 108(53.6) - 57(57.8) - 51(53.3)$$

$$= 5788.8 - 3294.6 - 2718.3 = -224.10$$

$$\sum k_i^2/r_i = \frac{108^2}{54} + \frac{(-57)^2}{57} + \frac{(-51)^2}{51} = 4(54) + 57 + 51 = \text{[redacted]}\ 324$$

$$SSC = \frac{(-224.10)^2}{324} = \text{[redacted]}\ 155.0025$$

- $d = \sum d_i \bar{Y}_{i.} = 0\bar{Y}_{1.} + \bar{Y}_{2.} - \bar{Y}_{3.} = 57.8 - 53.3 = 4.5$

$$\sum \frac{d_i^2}{r_i} = \frac{0}{54} + \frac{1}{57} + \frac{1}{51} = .037151703$$

$$SSD = \frac{(4.5)^2}{.037151703} = 545.0625$$

$$SSC + SSD = \text{[redacted]}\ 155.0025 + 545.0625 = \text{[redacted]}\ 700.065 \approx \text{[redacted]}\ SST$$

([redacted])

Response Curves for QUANTITATIVE TREATMENTS

- Sometimes treatment levels are quantitative (e.g. dose of drug, amount or intensity of exposure to stimulus, density of experimental units).
- When the treatments are quantitative, it may be that a polynomial response function can describe the relationship between $E[Y]$ & x (treatment).
- The treatment sum of squares can be partitioned into $t-1$ orthogonal polynomials using coefficients provided in Table XI - p. 623 for various values of t. See plots in Kuehl (p. 84) for linear, quadratic, & cubic trends

EXAMPLE - Dose response studies in laboratory animals w/ $t=10$ doses (0 to 4.5 by 0.5) and $r=6$ animals / dose.
We will fit the first 6 polynomials (orthogonal) as that is the highest order Kuehl gives coefficients for, and these account for 99% of the treatment sum of squares.

Source: Stephen J. Ruberg (1995): "Dose Response Studies. II. Analysis and Interpretation" Journal of Biopharmaceutical Statistics, 5(1), 15-42

(3.12) P. 3.14 - 3.16

DOSE	r	$\bar{y}_{i.}$	s_i	$\Sigma(\bar{y}_{i.}-\bar{y}_{..})^2$	s_i^2	$\hat{y}_i$	$s_{\hat{y}_i}$	95% CI for μ_i
0.0	6	25.5	2.6	632.02	6.76	25.33	3.17	(18.96, 31.6
0.5	6	23.9	4.0	715.03	16.00	24.47	3.01	(18.41, 30.5
1.0	6	27.7	3.3	526.24	10.89	27.30	2.42	(22.44, 32.
1.5	6	33.4	2.3	297.22	5.29	32.46	2.29	(27.86, 37.06,
2.0	6	40.5	10.5	102.82	110.25	42.71	2.23	(38.22, 47.19)
2.5	6	57.9	9.9	52.71	98.01	57.61	2.23	(53.12, 62.09)
3.0	6	74.4	14.6	564.54	213.16	71.48	2.29	(66.88, 76.08)
3.5	6	73.4	7.6	518.02	57.76	76.84	2.42	(71.78, 81.50)
4.0	6	73.5	4.5	522.58	20.25	71.93	3.01	(65.87, 77.98)
4.5	6	76.2	7.9	653.31	62.41	76.51	3.17	(70.14, 82.87)
		$\bar{y}_{..} = 50.64$		$\Sigma = 4584.49$	$\Sigma = 600.78$			

$SST = 6(4584.49) = 27506.94$ $\quad df_T = 10-1 = 9$

$SSE = (6-1)(600.78) = 3003.90$ $\quad df_E = 60-10 = 50$

$$F_0 = \frac{MST}{MSE} = \frac{27506.94/9}{3003.90/50} = \frac{3056.33}{60.08} = 50.87$$

$F_{.05,9,50} = 2.08$

22-141 50 SHEETS
22-142 100 SHEETS
22-144 200 SHEETS

Orthogonal Polynomial Coefficients (P_{ci}) From Table XI – P. 623 ($t=10$)

DOSE	$\bar{Y}_{i.}$	MEAN	LINEAR P_{1i}	QUADRATIC P_{2i}	CUBIC P_{3i}	QUARTIC P_{4i}	QUINTIC P_{5i}	SEXTIC P_{6i}
0.0	25.5	1	-9	6	-42	18	-6	3
0.5	23.9	1	-7	2	14	-22	14	-11
1.0	27.7	1	-5	-1	35	-17	-1	10
1.5	33.4	1	-3	-3	31	3	-11	6
2.0	40.5	1	-1	-4	12	18	-6	-8
2.5	57.9	1	1	-4	-12	18	6	-8
3.0	74.4	1	3	-3	-31	3	11	6
3.5	73.4	1	5	-1	-35	-17	1	10
4.0	73.5	1	7	2	-14	-22	-14	-11
4.5	76.2	1	9	6	42	18	6	3
λ		—	2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{5}{12}$	$\frac{1}{10}$	$\frac{11}{240}$
$\text{Sum} = \sum P_{ci}\bar{Y}_{i.}$		506.40	1172.40	(-13.10) -1[illegible].10	-1644.30	63.70	210.90	104.30
$\text{Divisor} = \sum P_{ci}^2$		10	330	132	8580	2860	780	660
$SSP_c = \frac{r(\sum P_{ci}\bar{Y}_{i.})^2}{\sum P_{ci}^2}$		—	24991.30	7.80	1890.72	8.51	342.14	98.90
$\hat{\alpha}_c = \frac{\sum P_{ci}\bar{Y}_{i.}}{\sum P_{ci}^2}$		50.64	3.55	-0.10	-0.19	0.022	0.27	0.16
$s^2_{\hat{\alpha}_c} = \frac{s^2}{r\sum P_{ci}^2}$		1.00	.0303	.0759	.0012	.0035	.0128	.0152

Analysis of Variance (up to ~~Quartic~~ SEXTIC TERM)

SOURCE	df	SS	MS	F	$F_{.05,9,50}$
DOSE	9	27506.94	3056.33	50.87	2.08
ERROR	50	3003.90	60.08		

Contrast	df	Contrast SS	MS	F	$F_{.05,1,50}$
LINEAR	1	24991.30	24991.30	415.97	4.04
QUADRATIC	1	7.80	7.80	0.13	4.04
CUBIC	1	1890.72	1890.72	31.47	4.04
QUARTIC	1	8.51	8.51	0.14	4.04
QUINTIC	1	342.14	342.14	5.69	4.04
SEXTIC	1	98.90	98.90	1.65	4.04

All other terms sum to 27506.94 − 27240.47 = 266.47,
So it is highly unlikely any high order terms are significant. Why?

Obtaining predicted values for the 10 doses:

$$\hat{Y}_i = \bar{Y}_{\bullet\bullet} + \sum_{j=1}^{6} \hat{\alpha}_j P_{ji} \qquad i = 1, \ldots, 10$$

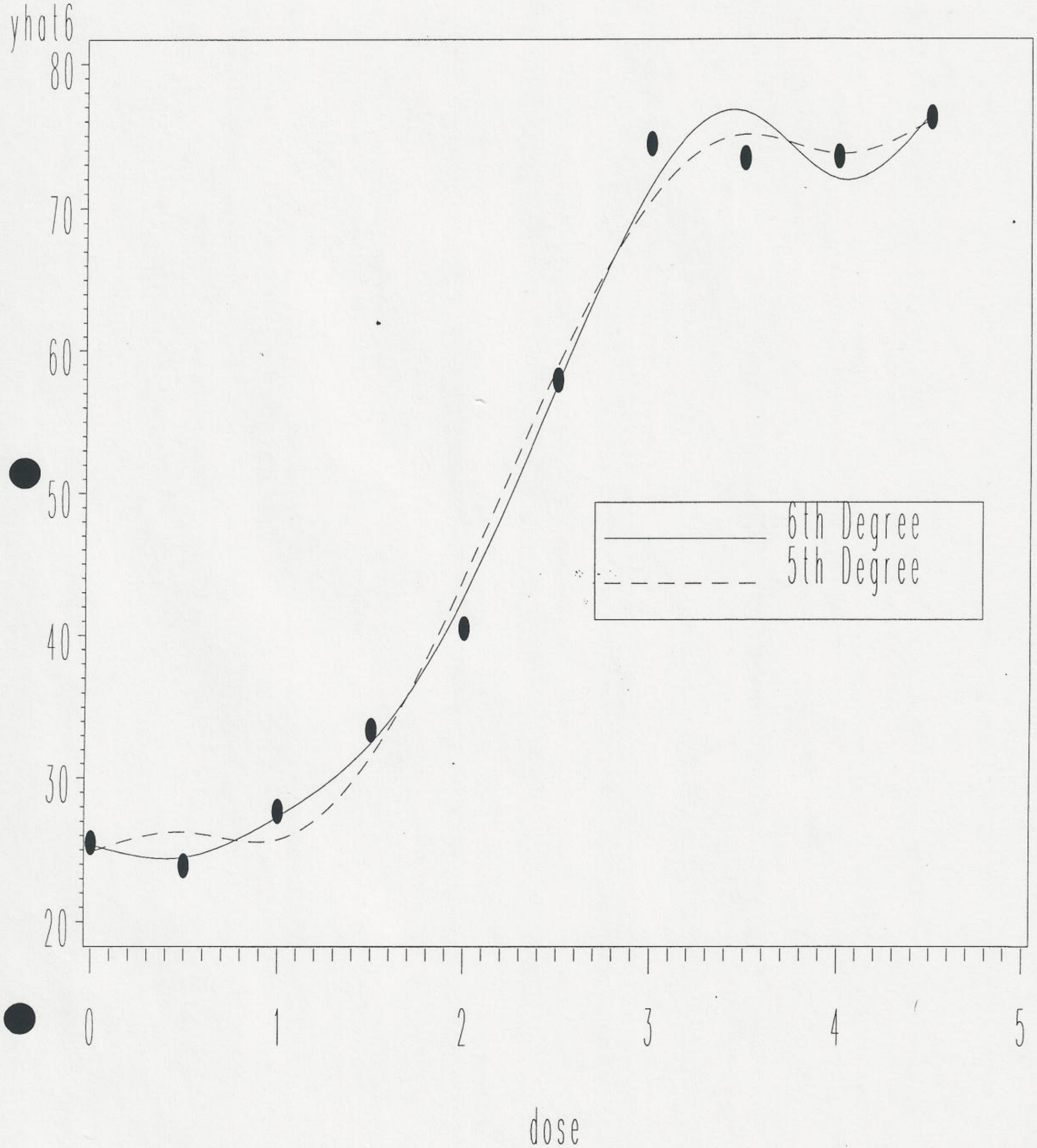
For dose 0: $\hat{Y}_0 = 50.64 + 3.55(-9) + (-0.10)(6) + (-0.19)(-42)$
$+ 0.022(18) + 0.27(-6) + 0.16(3)$

$= 50.64 - 31.95 - 0.60 + 7.98 + 0.40 - 1.62 + 0.48 = 25.33$

(Fitted Values for all doses are given on P. 3.17)

Dose Response Study - Orthogonal Polynomial Contrasts

3.15



$$Var(\hat{\alpha}_c) = Var\left(\frac{\sum P_{ci}\bar{y}_{i.}}{\sum P_{ci}^2}\right) = \left[\frac{1}{\sum P_{ci}^2}\right]^2 \sum P_{ci}^2 \, V(\bar{y}_{i.})$$

$$= \left[\frac{1}{\sum P_{ci}^2}\right]^{\cancel{2}} \cancel{\left[\sum P_{ci}^2\right]} \frac{\sigma^2}{r} = \frac{\sigma^2}{r\sum_{i=1}^{\wedge} P_{ci}^2}$$

$$\Rightarrow s^2_{\hat{\alpha}_c} = \frac{s^2}{r\sum_{i=1}^{\wedge} P_{ci}^2}$$

$$s^2_{\hat{y}_i} = s^2_{\bar{y}_{..}} + \sum_{j=1}^{6} P_{ji}^2 \, s^2_{\hat{\alpha}_j} \qquad i=1,\dots,10$$

Don't need to worry about covariances since these are orthogonal polynomials

NOTE: $s^2 = MSE = 60.08 \qquad r=6 \qquad s^2_{\bar{y}_{..}} = \frac{s^2}{rt} = \frac{60.08}{60} \approx 1.0$

For dose 0:

$$s^2_{\hat{y}_0} = s^2_{\bar{y}_{..}} + P_{1i}^2 s^2_{\hat{\alpha}_1} + P_{2i}^2 s^2_{\hat{\alpha}_2} + P_{3i}^2 s^2_{\hat{\alpha}_3} + P_{4i}^2 s^2_{\hat{\alpha}_4} + P_{5i}^2 s^2_{\hat{\alpha}_5} + P_{6i}^2 s^2_{\hat{\alpha}_6}$$

$$= 1.00 + (-9)^2(.0303) + 6^2(.0759) + (-42)^2(.0012) + 18^2(.0035) + (-6)^2(.0128) + 3^2(.0152) =$$

$$= 1.00 + 2.45 + 2.73 + 2.12 + 1.13 + 0.46 + 0.14 = 10.03 \Rightarrow s_{\hat{y}_0} = 3.17$$

(STD. ERRORS FOR ALL FITTED VALUES ARE GIVEN ON TABLE ON P. 3.12)

3.4 Multiple Comparisons Affect Error Rates

- $\alpha_C \equiv$ Comparisonwise error rate for a single comparison
- With t trt means, there are $\binom{t}{2} = \frac{t(t-1)}{2}$ possible comparisons. We can make up to $\binom{t}{2}$ type I errors if all t means are equal.
- $\alpha_E \equiv$ experimentwise error rate $\equiv$ accumulated risks for the individual comparisons.

Obtaining the maximum Error RATE

$n = \binom{t}{2}$ tests of the form statistics $t_0 = \dfrac{\bar{Y}_i - \bar{Y}_j}{\sqrt{s^2\left[\frac{1}{r_i} + \frac{1}{r_j}\right]}}$

(NOTE: These tests are not independent since they share common variance estimate and sample means).

- Upper limit for Type I error rate (experimentwise) can be obtained assuming independence.
- Assume $H_0: \mu_i - \mu_j = 0$ is true $\forall (i,j)$
  - $Pr\{\text{Type I error}\} = \alpha_c \Rightarrow Pr\{\text{correct decision}\} = 1-\alpha_c$
  - Let $X \equiv$ # of type I errors: $X \sim Bin(n, \alpha_c)$

$\Rightarrow p(x) = Pr\{X = x\} = \binom{n}{x} \alpha_c^{\,x}(1-\alpha_c)^{n-x} \quad x = 0, 1, \ldots, n$

$\Rightarrow Pr\{\text{No Type I errors}\} = p(0) = (1-\alpha_c)^n$

$\Rightarrow Pr\{\text{@ Least 1 Type I error}\} = 1-(1-\alpha_c)^n = \alpha_E$

$\Rightarrow$ Solving for α_c that gives prescribed α_E: $\boxed{\alpha_c = 1-(1-\alpha_E)^{1/n}}$

See Table on P.93 of Kuehl.

2 Types of Exprimetwise Error Rate

① Weak Sense - Defined under configuration $\mu_1 = \cdots = \mu_t$

② Strong Sense - Probability of @ least one wrong decision over all parameter configurations.

STRONG: $$\text{Error RATE} = \sup_{\{\mu_i\}} P_{\{\mu_i\}}\left[\text{@ least one incorrect assertion}\right].$$

WEAK ≡ STRONG WHEN SUPREMUM occurs @ $\mu_1 = \mu_2 = \cdots = \mu_t$

3.5 Simultaneous Statistical Inference

Types of Contrasts Constructed

- Planed Contrasts among treatment means
- Orthogonal polynomial contrasts
- Multiple Comparisons with the best treatment
- " " " " Control "
- All pairwise comparisons

Strength of Infrence (each w/ given levels of confidence)

STRONG

- Simultaneous Confidence Intervals => direction & magnitude.
- Confident Directions => Direction only ($C_i > 0$, $C_i < 0$)
- Confident Inequalities => Inequality only ($C_i \neq 0$)
- Individual Comparisons => Ignore Simultaneity of inference

WEAK

We will focus on simultaneous 2-sided CI's, the strongest inference.

Bonferroni's Inequality (Very general approach to simultaneous inference)

- $\alpha_E \leq \sum_{i=1}^{\hat{n}} \alpha_{c_i} = n\alpha_c$ when there are n comparisons, each conducted @ level $\alpha = \alpha_c$

- $\Rightarrow$ For a chosen experimentwise error rate α_E, conduct each @ $\alpha_c = \frac{\alpha_E}{n}$ (conservative due to inequality above)

n	α_E	α_c
3	.05	.0167
3	.01	.0033
4	.05	.0125
4	.01	.0025
	⋮	
10	.05	.0050
10	.01	.0010

Table V pp. 595-6

Give $t_{\frac{\alpha}{2}, k, \nu}$ ~~for~~,
$\alpha = .01, .05$
k = # of comparisons (CI's)
ν = Error df (degrees of freedom associated w/ s^2).

Algorithm

① Identify k contrasts of interest: $C_i = \sum_{j=1}^{t} c_{ij}\mu_j \quad i = 1, \ldots, k$

② Compute k estimates: $c_i = \sum_{j=1}^{t} c_{ij}\bar{Y}_{j\cdot} \quad i = 1, \ldots, k$

③ Compute k std errors: $\quad i = 1, \ldots, k$

$$S_{c_i} = \sqrt{s^2 \sum_{j=1}^{t} \frac{c_{ij}^2}{r_j}} \qquad i = 1, \ldots, k$$

④ Compute k simultaneous $(1-\alpha_E)100\%$ Confidence Intervals:

$$\text{CI for } C_i: \quad c_i \pm t_{\frac{\alpha_E}{2}, k, \nu} \cdot S_{c_i} \qquad i = 1, \ldots, k$$

Scheffe's Test for all possible Comparisons (Contrasts)

- Good for all possible linear contrasts of $\{\mu_i\}$ simultaneously with α_E

Algorithm

① Identify k contrasts of interest: $C_i = \sum_{j=1}^{t} c_{ij}\mu_j \quad i=1,\ldots$

② Compute k estimates: $c_i = \sum_{j=1}^{t} c_{ij}\,\bar{y}_{j.} \qquad i=1,\ldots,k$

③ Compute k STD. ERRORS:

$$S_{c_i} = \sqrt{S^2 \sum_{j=1}^{t} \frac{c_{ij}^2}{r_j}} \qquad i=1,\ldots,k$$

④ Compute k Simultaneous $100(1-\alpha_E)\%$ CI's

$$c_i \pm \sqrt{(t-1)\,F_{\alpha_E,t-1,\nu}} \cdot S_{c_i} \qquad i=1,\ldots,k$$

3.6 Multiple Comparisons w/ the Best TRT

- Goal: To Determine the best treatment (or a subset of treatments that contain the best).
- Assume that high scores are good (obvious adjustment for case where low scores are good).

Constrained MCB ALGORITHM (Interval Must Contain 0)

(Assumes no two trts have identical means)

① Identify parameters of interest: $\mu_i - \max_{j \neq i} \mu_j \quad i = 1, \ldots, t$

$\mu_i - \max_{j \neq i} \mu_j > 0 \Rightarrow$ TRT i best

$\mu_i - \max_{j \neq i} \mu_j < 0 \Rightarrow$ TRT i not best

② Compute sample analogues to ① $\quad D_i = \bar{Y}_{i.} - \max_{j \neq i} \bar{Y}_{j.}$

$i = 1, \ldots, t$

③ Compute critical quantity $(r_1 = r_2 = \cdots = r_t = r)$

$$M = d_{\alpha_E, k, \nu} \sqrt{\frac{2s^2}{r}}$$

$\boxed{k = t \text{ comparisons} \quad \nu \equiv \text{Error df } (t(r-1))}$

$d_{\alpha_E, k, \nu}$ values given in Table VI pp. 597-600

for $\alpha_E = .05, .01$ & 1- & 2-sided tests.

$\boxed{\text{Use 1-sided}}$

④ Lower Bound: $\quad L = \begin{cases} D_i - M & \text{if } D_i - M < 0 \\ 0 & \text{o.w.} \end{cases}$

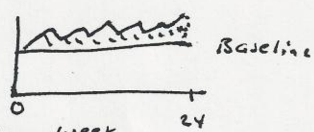
Upper Bound: $\quad U = \begin{cases} D_i + M & \text{if } D_i + M > 0 \\ 0 & \text{o.w.} \end{cases}$

(5) If interval includes 0 or has a lower bound of 0, μ_i is one of highest (trt i is one of the best). If the upper bound of interval is 0, TRT i is not one of the best.

EXAMPLE – 3 TRTS FOR HIV

Collier, et al) (1996): NEJM, 334: 1011–1017

- RESPONSE: Y = AREA UNDER LN(CD4*) COUNT VS. TIME CURVE (stores can be +/−)



- t = 3 TREATMENTS

TRT 1: SAQUINAVIR, ZIDOVUDINE, ZALCITABINE (SZZ)
TRT 2: SAQ, ZID (SZ) TRT 3: ZID, ZAL (ZZ)

- r = 90 reps / treatment

TRT	r_i	$\bar{Y}_{i.}$	S_i	$r_i(\bar{Y}_{i.}-\bar{Y}_{..})^2$	$(r_i-1)S_i^2$	D_i
SZZ	90	12.2	18.97	3841.60	32027.62	7.1
SZ	90	5.1	19.92	28.90	35315.77	−7.1
ZZ	90	−0.3	20.87	3204.10	38764.56	−12.5
		$\bar{Y}_{..} = 5.67$		SST = 7074.60	SSE = 106107.95	

$\alpha_E = .05$

$S^2 = \frac{SSE}{270-3} = 397.41$

$d_{.05,3,267} \approx 2.07$ (interpolating between $\nu = 120, \infty$)

$M = 2.07\sqrt{2(397.41)/90} = 6.15$

i	$L(\mu_i - \max_{j \neq i} \mu_j)$	$U(\cdot)$
1	0 (since $7.1 - 6.15 > 0$)	$7.1 + 6.15 = 13.25$
2	$-7.1 - 6.15 = -13.25$	0 (since $-7.1 + 6.15 < 0$)
3	$-12.5 - 6.15 = -18.65$	0 ($-12.5 + 6.15 < 0$)

Conclude SZZ is the unique Best TRT

Multiple Comparisons with the smallest Mean

- Parameter: $\mu_i - \min_{j \neq i} \mu_j \qquad i = 1, \dots, t$

- Estimate: $D_i = \bar{y}_{i.} - \min_{j \neq i} \bar{y}_{j.} \qquad i = 1, \dots, t$

- $L = \begin{cases} D_i - M & \text{if } D_i - M < 0 \\ 0 & \text{o.w} \end{cases}$ (M as above)

  $U = \begin{cases} D_i + M & \text{if } D_i + M > 0 \\ 0 & \text{o.w.} \end{cases}$

- Decision Rule: If interval includes 0 or has 0 as an upper bound TRT i is one of best. If interval has 0 for lower bound, trt i is not one of best.

$\boxed{\text{See Kuehl Example w/ Meat Storage pp 102-103}}$

3.7 Comparison of All TRTS TO A CONTROL

- Goal: Compare t-1 "active" treatments to a Control.
- Can make use of 1-sided (direction chosen prior to observing data) or 2-sided simultaneous confidence Intervals.
- Often helpful to identify control group as group t or group 1 (or generic group c).

Algorithm

- Parameter: $\mu_i - \mu_c$ $\quad i = 1, \dots, t-1$
- Estimate: $\bar{y}_{i\cdot} - \bar{y}_{c\cdot}$ $\quad i = 1, \dots, t-1$
- Standard Error: $S_{\bar{y}_{i\cdot} - \bar{y}_{c\cdot}} = \sqrt{\frac{2S^2}{r}}$ $\quad (r_1 = r_2 = \dots = r_t)$
- Table value $d_{\alpha_E, k, \nu}$ (Dunnett's Table) Table VI pp. 597-600

$\alpha_E = .05, .01$
$k = \#$ of comparisons $= t-1$
$\nu =$ df ERROR $= t(r-1)$
1-sided & 2-sided

2-Sided Simultaneous $(1-\alpha_E)100\%$ CI's for $\mu_i - \mu_c$

$$(\bar{y}_{i\cdot} - \bar{y}_{c\cdot}) \pm \underbrace{d_{\alpha_E, k, \nu}}_{\text{2-sided}} \cdot S_{\bar{y}_{i\cdot} - \bar{y}_{c\cdot}} \qquad i = 1, \dots, t-$$

Determine differences wrt control by range of values in interval

1-sided (Lower Bound) $(1-\alpha_E)100\%$ CI's for $\mu_i -$

$$\left[(\bar{y}_{i\cdot} - \bar{y}_{c\cdot}) - \underbrace{d_{\alpha_E, k, \nu}}_{\text{1-sided}} \cdot S_{\bar{y}_{i\cdot} - \bar{y}_{c\cdot}}, \ \infty \right)$$

1-sided (Upper bound) $(1-\alpha)100\%$ CI's for $\mu_i - \mu_c$

$$\left(-\infty, \ (\bar{y}_{i\cdot} - \bar{y}_{c\cdot}) + \underbrace{d_{\alpha_E, k, \nu}}_{\text{1-sided}} \cdot S_{\bar{y}_{i\cdot} - \bar{y}_{c\cdot}} \right]$$

Juskevich & Guyer (1990) Science 249:875-884 — 3.25

EXAMPLE: BOVINE GROWTH HORMONE: HUMAN FOOD SAFETY EVALUATION

GOAL: STUDY EFFECT OF VARIOUS DOSES OF BGH ON Body weight change (g) on male rats (85 days)

TRT			
1	Control:	0 rbGH (mg/kg per day)	
2	Subcutaneous:	1 rbgH	injected
3	oral:	0.1 rbgH	oral
4	"	0.5	
5	"	5.0	
6	"	50.0	

RESEARCH HYPOTHESES: (ALL versus TRT 1)

(A) Positive Effect for TRT 2 (Positive con

(B) No effects among TRTS 3-6

TRT	r_i	$\bar{y}_{i.}$	s_i	$r_i(\bar{y}_{i.}-\bar{y}_{..})^2$	$(r_i-1)s_i^2$	$\bar{y}_{i.}-\bar{y}_{c.}$
1	30	324	39.2	10083.3	44562.56	—
2	30	432	60.3	241203.3	105446.61	108
3	30	327	39.1	7053.3	44335.49	3
4	30	318	53.0	17763.3	81461.00	−6
5	30	325	46.3	9013.3	62167.01	1
6	30	328	43.0	6163.3	53621.00	4
		$\bar{y}_{..} = 342.33$		SST=291280	SSE=391593.67	

$$s^2 = \frac{SSE}{174} = 2250.54$$

2-Sided simultaneous 95% CIs

$K = t-1 = 6-1=5 \qquad \nu = t(r-1) = 6(30-1) = 174 \qquad d_{.05,5,174} \simeq 2.55$ (2-sided)

PIVOTAL QUANTITY: $d_{\alpha E,k,\nu} \cdot \sqrt{\frac{2s^2}{r}} = 2.55\sqrt{\frac{2(2250.54)}{30}} = 31.23$

TRT	$\bar{y}_{i.}-\bar{y}_{c.}$	95% SCI	Different from control?
1	—	—	—
2	108	(76.77, 139.23)	YES
3	3	(−28.23, 34.23)	NO
4	−6	(−37.23, 25.23)	NO
5	1	(−30.23, 32.23)	NO
6	4	(−27.23, 35.23)	NO

3.8 Pairwise Comparison of All Treatments

• GOAL: COMPARE ALL PAIRS OF TRT MEANS

$$\binom{t}{2} = \frac{t(t-1)}{2} \quad \text{PAIRS (COMPARISONS)}$$

① Bonferroni's Approach

Choose $K = \frac{t(t-1)}{2}$ & Apply Bonferroni's Method described previously

$$(\bar{Y}_{i.} - \bar{Y}_{j.}) \pm t_{\frac{\alpha_E}{2}, K, \nu} \sqrt{\frac{S_i^2}{r_i} + \frac{S_j^2}{r_j}} \quad \forall\ i < j$$

② Tukey's Honestly Significant Difference

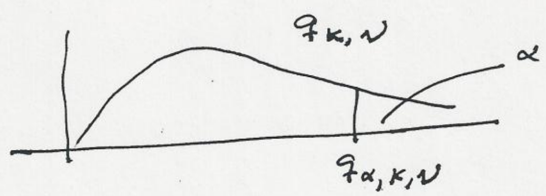
Studentized Range Distribution

Suppose $Y_1, \ldots, Y_n \sim NID(\mu, \sigma^2)$

Then $\frac{\max Y_i - \min Y_i}{S} \sim q_{n,\nu}$

where $n \equiv$ sample size, $\nu \equiv$ degrees of freedom for S

Critical values of q are given in Table VII, pp. 601-2 for tail areas $\alpha = .05, .01$



Application to multiple Comparisons

$\bar{Y}_{i.} \sim N(\mu_i, \frac{\sigma^2}{r})$ $\quad i = 1, \dots, t$ (Assuming $r_1 = r_2 = \dots = r_t = r$)

(independent in CRD)

$\bar{Y}_{i.} - \mu_i \sim NID(0, \frac{\sigma^2}{r})$ $\quad s^2 = MSE$ is estimate of σ^2 w/ $\nu = t(r-1)$ df

$$\Rightarrow \frac{\max(\bar{Y}_{i.} - \mu_i) - \min(\bar{Y}_{i.} - \mu_i)}{s/\sqrt{r}} \sim q_{t,\nu}$$

$$\Rightarrow Pr\left\{\max(\bar{Y}_{i.} - \mu_i) - \min(\bar{Y}_{i.} - \mu_i) \geq q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}}\right\} = \alpha$$

$$\Rightarrow Pr\left\{|(\bar{Y}_{i.} - \mu_i) - (\bar{Y}_{j.} - \mu_j)| \leq q_{\alpha,t,\nu} \cdot \sqrt{s^2/r}\right\} = 1-\alpha \quad \forall i,j$$

(Since the equality holds for largest difference)

$$\Rightarrow Pr\left\{|(\bar{Y}_{i.} - \bar{Y}_{j.}) - (\mu_i - \mu_j)| \leq q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}}\right\} = 1-\alpha \quad \forall i,j.$$

$$\Rightarrow Pr\left\{-q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}} \leq (\bar{Y}_{i.} - \bar{Y}_{j.}) - (\mu_i - \mu_j) \leq q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}}\right\} = 1- \quad \forall i,j$$

$$\Rightarrow Pr\left\{-(\bar{Y}_{i.} - \bar{Y}_{j.}) - q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}} \leq -(\mu_i - \mu_j) \leq -(\bar{Y}_{i.} - \bar{Y}_{j.}) + q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}}\right\} = 1- \quad \forall i,j$$

$$\Rightarrow Pr\left\{(\bar{Y}_{i.} - \bar{Y}_{j.}) + q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}} \geq \mu_i - \mu_j \geq (\bar{Y}_{i.} - \bar{Y}_{j.}) - q_{\alpha,t,\nu} \sqrt{\frac{s^2}{r}}\right\} = 1-\alpha$$

(1−α)100% Simultaneous CI's for $\mu_i - \mu_j$ (Tukey's HSD)

$$(\bar{Y}_{i.} - \bar{Y}_{j.}) \pm q_{\alpha,t,\nu} \cdot \sqrt{\frac{s^2}{r}} \quad \forall i,j$$

NOTES: ① Does not use $\frac{\alpha}{2}$ from Studentized Range Distribution

② Does not use $s.e.(\bar{Y}_{i.} - \bar{Y}_{j.})$ explicitly.

Tests of Homogeneity : $\mu_1 = \mu_2 = \cdots = \mu_t$

Tests set up with experimentwise error rates based on homogeneity (Note: Tukey's and Bonferroni's approaches are <u>not</u> set up under this assumption).

Least Significant Differnce (LSD)

① Parameters: $\mu_i - \mu_j \quad (i>j)$

② ESTIMATES: $\bar{Y}_{i.} - \bar{Y}_{j.} \quad i>j$

③ EST. STD. ERROR: $\sqrt{s^2\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$

④ $LSD(\alpha) = t_{\frac{\alpha}{2},\nu} \cdot \sqrt{s^2\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$

⑤ ~~Sim~~ Confidence Intervals

$$(\bar{Y}_{i.} - \bar{Y}_{j.}) \pm LSD(\alpha) \equiv (\bar{Y}_{i.} - \bar{Y}_{j.}) \pm t_{\frac{\alpha}{2},\nu} \cdot \sqrt{s^2\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$$

⑥ Tests: Conclude $H_A: \mu_i \neq \mu_j$ if

$$|\bar{Y}_{i.} - \bar{Y}_{j.}| > LSD(\alpha)$$

This test will tend to reject too often for large values of t.

Protected LSD (Controls weak sense experimentwise error rate)

① Conduct the F-test of $H_0: \mu_1 = \cdots = \mu_t \quad H_A:$ Not all μ_i are =

IF REJECT – CONDUCT LSD on all pairs

IF FAIL TO REJECT – STOP (Conclude no pairs differ)

Student-Newman-Keuls (SNK) Multiple RANGE TEST

STUDENTIZED RANGE TEST THAT PROVIDES A HOMOGENEITY TEST IN WEAK SENSE.

ALGORITHM

1. ORDER MEANS $\bar{y}_{(1)} \le \bar{y}_{(2)} \le \cdots \le \bar{y}_{(t)}$

2. COMPUTE $SNK(k, \alpha_E) = q_{\alpha, k, \nu} \sqrt{\frac{s^2}{r}}$

   where $k \equiv$ # of means in range $k = 2, 3, \ldots, t$
   $\nu \equiv$ Error df $= t(r-1)$

3. Conclude $H_A: \mu_i \ne \mu_j$ (vs. $H_0: \mu_i = \mu_j$) if

   $|\bar{y}_{i.} - \bar{y}_{j.}| \ge SNK(k, \alpha_E)$ where k is the number of means in range containing i, j

|Means diff|	K
$\|\bar{y}_{(1)} - \bar{y}_{(t)}\|$	t
$\|\bar{y}_{(1)} - \bar{y}_{(t-1)}\|$, $\|\bar{y}_{(2)} - \bar{y}_{(t)}\|$,	$t-1$
$\vdots$	

NOTE: All procedures that use a critical value for $|\bar{y}_{i.} - \bar{y}_{j.}|$ could be used to form confidence intervals and obtain identical conclusions.