Chapter 8 - Complete Block Designs

BLOCKS - GROUPS OF HOMOGENEOUS EXPERIMENTAL UNITS

- PROXIMITY
- Physical characteristics
- TIME
- MANAGEMENT OF TASKS
- INDIVIDUALS (TECHNICIANS / SUBJECTS)

BLOCKS ARE ALMOST ALWAYS RANDOM

Model (Equation 8.14, p. 275)

$$Y_{ij} = \mu + \tau_i + b_j + e_{ij} \qquad i=1,\dots,t \quad j=1,\dots,r \text{ (\# of blocks)}$$

$$\sum_i \tau_i = 0 \qquad b_j \sim NID(0, \sigma_b^2) \qquad e_{ij} \sim NID(0, \sigma^2)$$

TREATMENTS: $SSA = r \sum_{i=1}^{t} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 \qquad df_A = t-1$

BLOCKS: $SSB = t \sum_{j=1}^{r} (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2 \qquad df_B = r-1$

ERROR: $SSE = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 \qquad df_E = (t-1)(r-1)$

$$E[Y_{ij}] = \mu + \tau_i + 0 + 0 = \mu + \tau_i$$

$$Cov(Y_{ij}, Y_{i'j'}) = Cov(\mu + \tau_i + b_j + e_{ij}, \mu + \tau_{i'} + b_{j'} + e_{i'j'})$$

$$= Cov(b_j + e_{ij}, b_{j'} + e_{i'j'}) = \begin{cases} \sigma^2 + \sigma_b^2 & i=i', j=j' \\ \sigma_b^2 & i \neq i', j=j' \\ 0 & j \neq j' \end{cases}$$

$$E[\bar{y}_{i.}] = E\left[\frac{1}{r}\sum_j y_{ij}\right] = \frac{1}{r}\sum_j E(y_{ij})$$

$$= \frac{1}{r}\left[r(\mu+\tau_i)\right] = \mu+\tau_i$$

$$V[\bar{y}_{i.}] = V\left[\frac{1}{r}\sum_{j=1}^{r} y_{ij}\right] = \left(\frac{1}{r}\right)^2 V\left[\sum_j y_{ij}\right]$$

$$= \left(\frac{1}{r}\right)^2 \left\{ \sum_{j=1}^{r} V(y_{ij}) + 2\sum\sum_{j<j'} \overbrace{\cancel{\operatorname{cov}(y_{ij}, y_{ij'})}}^{0} \right\}$$

$$= \frac{\sigma^2+\sigma_b^2}{r} \Rightarrow \boxed{E[\bar{y}_{i.}^2] = (\mu+\tau_i)^2 + \frac{1}{r}(\sigma^2+\sigma_b^2)}$$

$$E[\bar{y}_{..}] = E\left[\frac{1}{rt}\sum_i\sum_j y_{ij}\right] = \frac{1}{rt}\sum_i\sum_j E(y_{ij})$$

$$= \frac{1}{rt}\sum_i r(\mu+\tau_i) = \frac{1}{rt}\left[rt\mu + r\overbrace{\cancel{\sum_i \tau_i}}^{0}\right]$$

$$= \mu$$

$$\cancel{E}V[\bar{y}_{..}] = V\left[\frac{1}{rt}\sum_i\sum_j y_{ij}\right] = \left(\frac{1}{rt}\right)^2 V\left[\sum_i\sum_j y_{ij}\right]$$

$$= \left(\frac{1}{rt}\right)^2 \left\{ \sum_i\sum_j V(y_{ij}) + 2\sum\sum_{i<i'}\sum_j \operatorname{cov}(y_{ij}, y_{i'j}) \right.$$

$$\left. + 2\sum_i \sum\sum_{j<j'} \operatorname{cov}(y_{ij}, y_{ij'}) + 4\sum\sum_{i<i'}\sum\sum_{j<j'} \operatorname{cov}(y_{ij}, y_{i'j'}) \right\}$$

$$= \left(\frac{1}{rt}\right)^2 \left\{ rt(\sigma_b^2+\sigma^2) + rt(t-1)\sigma_b^2 + 0 + 0 \right\}$$

$$= \frac{1}{rt}\left\{ \sigma_b^2 + \sigma^2 + (t-1)\sigma_b^2 \right\} = \frac{1}{rt}\left\{ \sigma^2 + t\sigma_b^2 \right\} = V(\bar{y}_{..})$$

$$\Rightarrow \boxed{E[\bar{y}_{..}^2] = \mu^2 + \frac{1}{rt}\left\{\sigma^2 + t\sigma_b^2\right\}}$$

$$E[SSA] = E\left[r\sum_i \bar{y}_{i.}^2 - rt\,\bar{y}_{..}^2 \right]$$

$$= r\sum_i \left[(\mu+\tau_i)^2 + \frac{1}{r}(\sigma^2 + \sigma_b^2) \right] - rt\left\{ \mu^2 + \frac{1}{rt}(\sigma^2 + t\sigma_b^2) \right\}$$

$$= \left[\cancel{rt\mu^2} + 2r\mu \cancel{\sum \tau_i} + r\sum_i \tau_i^2 + t\sigma^2 + \cancel{t\sigma_b^2} \right]$$

$$- \left[\cancel{rt\mu^2} + \sigma^2 + \cancel{t\sigma_b^2} \right]$$

$$= r\sum_i \tau_i^2 + (t-1)\sigma^2 \Rightarrow \boxed{E(MSA) = \sigma^2 + \frac{r\sum_i \tau_i^2}{t-1}}$$

$H_0: \tau_1 = \cdots = \tau_t = 0 \quad \left(\sum_i \tau_i^2 = 0\right)$

H_A: NOT ALL $\tau_i = 0 \quad \left(\sum_i \tau_i^2 > 0\right)$

T.S. $F_A = \dfrac{MSA}{MSE}$

RR: $F_A \geq F_{\alpha, t-1, (t-1)(r-1)}$

$$V[\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}] = V[\bar{y}_{i\cdot}] + V[\bar{y}_{i'\cdot}] - 2\,Cov(\bar{y}_{i\cdot}, \bar{y}_{i'\cdot})$$

$$= 2\left(\frac{\sigma^2 + \sigma_b^2}{r}\right) - 2\,Cov\left(\frac{1}{r}\sum_j y_{ij}, \ \frac{1}{r}\sum_j y_{i'j}\right)$$

$$= 2\left(\frac{\sigma^2 + \sigma_b^2}{r}\right) - 2\left(\frac{1}{r^2}\right)\left[\sum_j Cov(y_{ij}, y_{i'j}) + \cancel{2\sum_{j<j'}\sum Cov(y_{ij}, y_{i'j'})}^{\,0}\right]$$

$$= 2\left(\frac{\sigma^2 + \sigma_b^2}{r}\right) - 2\left(\frac{1}{r^2}\right) r\,\sigma_b^2$$

$$= 2\left(\frac{\sigma^2 + \sigma_b^2}{r}\right) - 2\left(\frac{\sigma_b^2}{r}\right) = \frac{2\sigma^2}{r}$$

$$\Rightarrow \quad S_{\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}} = \sqrt{\frac{2\,MSE}{r}}$$

$(1-\alpha)100\%$ CI for $\mu_i - \mu_{i'}$

$$(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}) \pm t_{\frac{\alpha}{2}, (t-1)(r-1)} \sqrt{\frac{2\,MSE}{r}}$$

OBVIOUS ADJUSTMENTS FOR SIMULTANEOUS CONFIDENCE INTERVALS

BONFERRONI: $(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}) \pm t_{\frac{\alpha}{2}, K, (t-1)(r-1)} \sqrt{\frac{2\,MSE}{r}}$

Tukey: $(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot}) \pm q_{.05, K, (t-1)(r-1)} \sqrt{\frac{MSE}{r}}$

Gain in Efficiency From USING RBD vs. CRD

(1) Compute estimate of σ^2 for RCB: $S^2_{RCB} = MSE$

(2) ESTIMATE OF σ^2 from the CRD: $S^2_{CR} = \dfrac{SS\ Blocks + r(t-1)MSE}{rt-1}$

(3) Relative Efficiency (uncorrected for degrees of freedom):

$$RE = \frac{S^2_{CR}}{S^2_{RCB}}$$

(ignores degrees of freedom for estimates of σ^2)

(4) Correction Factor $\equiv \dfrac{(f_{rcB}+1)(f_{cr}+3)}{(f_{rcb}+3)(f_{cr}+1)}$

$f_{rcb} \equiv$ Error df for RCB: $(t-1)(r-1)$

$f_{cr} \equiv$ " " for CRD: $t(r-1)$

(5) Multiply correction factor by RE

Bachmann, et al, (1995) "Controlled study of the putative interaction between famotidine & theophylline in patients w/ chronic obstructive pulmonary disorder" Journal of Clinical Pharmacology 35: 529-535. ~~8~~ 6

Example 6.3 In Example 1.5, we plotted data from a study quantifying the interaction between theophylline and two drugs (famotidine and cimetidine) in a three–period crossover study that cluded receiving theophylline with a placebo control (Bachmann, et al.,1995). We would like to compare the mean theophylline clearances when it is taken with each of the three drugs: cimetidine, famotidine, and placebo. Recall from Figure 1.5 that there was a large amount of subject–to–subject variation. In the RBD, we control for that variation when comparing the three treatments. The raw data, as well as treatment and subject (block) means are given in Table 6.7. The Analysis of Variance is given in Table 6.8. Note that in this example, we are comparing $k = 3$ treatments in $b = 14$ blocks.

We can now test for treatment effects, and if necessary use Tukey's method to make pairwise comparisons among the three drugs ($\alpha = 0.05$ significance level).

1. $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0 \quad (\mu_1 = \mu_2 = \mu_3)$ (No drug effect on theophylline clearance)
2. $H_A :$ Not all α_i are 0 (Drug effects exist)
3. T.S. $F_{obs} = \frac{MST}{MSE} = 10.64$
4. R.R.: $F_{obs} \geq F_{\alpha,k-1,(b-1)(k-1)} = F_{0.05,2,26} = 3.37$ (See P. 8.7)
5. p-value: $P(F \geq F_{obs}) = P(F \geq 10.64) = 0.0004$

Since we do reject H_0, and conclude differences exist among the treatment means, we will use Tukey's method to determine which drugs differ significantly. Recall that for Tukey's method, we compute simultaneous confidence intervals of the form given below, with k being the number of reatments (k=3), n the total number of observations ($n = bk$=3(14)=42), and n_i the number of ae rements per treatment ($n_i = b = 14$).

$$(\overline{y}_i - \overline{y}_j) \pm q_{\alpha,k,n-k}\sqrt{MSE(\frac{1}{n_i})} \implies (\overline{y}_i - \overline{y}_j) \pm 3.514\sqrt{0.33(\frac{1}{14})} \implies (\overline{y}_i - \overline{y}_j) \pm 0.54$$

he corresponding simultaneous 95% confidence intervals and conclusions are given in Table 6.9. e conclude that theophylline has a significantly lower clearance when taken with cimetidine than

Comparison	$\overline{y}_i - \overline{y}_j$	CI	Conclusion
Cimetidine vs Famotidine	$2.26 - 3.16 = -0.90$	$(-1.44, -.36)$	$C < F$
Cimetidine vs Placebo	$2.26 - 3.08 = -0.82$	$(-1.36, -.28)$	$C < P$
Famotidine vs Placebo	$3.16 - 3.08 = 0.08$	$(-0.46, 0.62)$	$F = P$

Table 6.9: Tukey's simultaneous 95% CI's for theophylline interaction data (RBD)

en taken with famotidine or placebo. No difference appear to exist when theophylline is taken h famotidine or with placebo. While cimetidine appears to interact with theophylline, famotidine s not appear to interact with it in patients with chronic obstructive pulmonary disease.

8.7

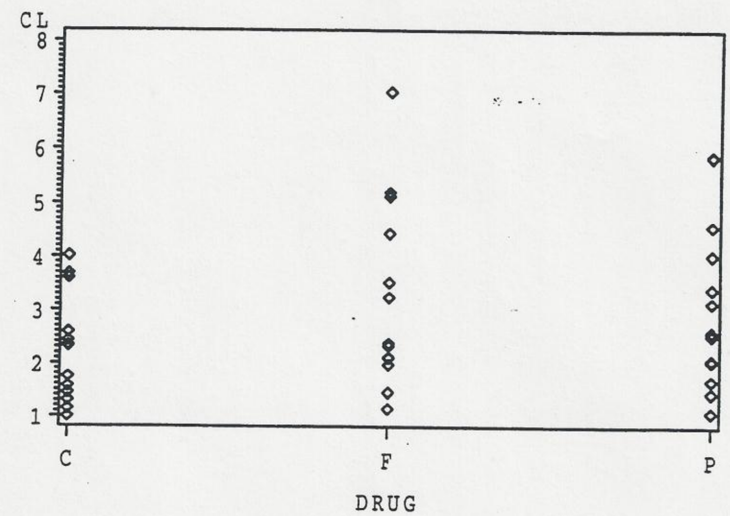


Figure 1.6: Plot of theophylline clearance vs interacting drug

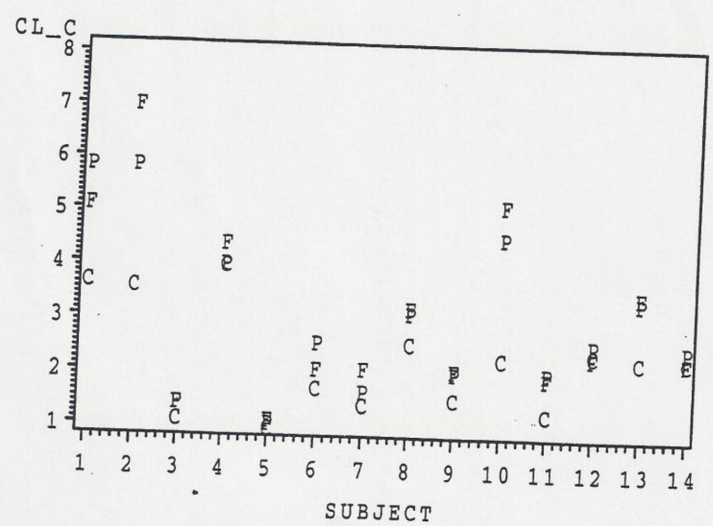


Figure 1.7: Plot of theophylline clearance vs subject with interacting drug as plotting symbol

8.8

Subject	Interacting Drug Cimetidine	Famotidine	Placebo	Subject Mean
1	3.69	5.13	5.88	4.90
2	3.61	7.04	5.89	5.51
3	1.15	1.46	1.46	1.36
4	4.02	4.44	4.05	4.17
5	1.00	1.15	1.09	1.08
6	1.75	2.11	2.59	2.15
7	1.45	2.12	1.69	1.75
8	2.59	3.25	3.16	3.00
9	1.57	2.11	2.06	1.91
10	2.34	5.20	4.59	4.04
11	1.31	1.98	2.08	1.79
12	2.43	2.38	2.61	2.47
13	2.33	3.53	3.42	3.09
14	2.34	2.33	2.54	2.40
Trt Mean	2.26	3.16	3.08	2.83

Table 6.7: Theophylline clearances (liters/hour) when drug is taken with interacting drugs

ANOVA

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
TREATMENTS	7.01	2	3.51	10.64
BLOCKS	71.81	13	5.52	
ERROR	8.60	26	0.33	
TOTAL	87.42	41		

$t=3$

$r=14$

Table 6.8: Analysis of Variance table for theophylline interaction data (RBD)

$$RE = \frac{S^2_{CR}}{S^2_{RCB}} \qquad S^2_{CR} = \frac{71.81 + 3(14-1)(0.33)}{3(14)-1} = 1.44$$

$$S^2_{RCB} = 0.33 \quad \Rightarrow RE = \frac{1.44}{0.33} = 4.36 \qquad CF = \frac{(26+1)(28+3)}{(26+3)(28+1)} = .995$$

Latin Square Design (2 BLOCKING CRITERIA)

EXAMPLE 1 Field Trials on a SQUARE GRID

Column

Row	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

- A-E ≡ TRTS
- NOTE THAT EACH TRT APPEARS ONCE IN EACH ROW & COLUMN

EXAMPLE 2 - COMPARISON OF 4 TIRES ON 4 CARS

POSITION

CAR	L/F	L/R	R/F	R/R
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

- A ≡ D ≡ BRANDS (FIXED)
- CARS (RANDOM, PRESUMABLY)
- POSITIONS (FIXED)

Randomization Procedure

1. Select a $t \times t$ standard square @ random (Kuehl, pp. 307-9)
2. Randomly order all but the first row
3. Randomly order all columns
4. Randomly Assign Trts to letters

NOTATION

y_{ij} ≡ observation in i^{th} row & j^{th} Column $\quad i = 1, \ldots, t$; $j = 1, \ldots, t$

$\bar{y}_{i.}$ ≡ Mean for row i $\quad \bar{y}_{.j}$ ≡ mean for row j

$\bar{y}_k$ ≡ mean for trt k $\quad \bar{y}_{..}$ ≡ overall mean

ANALYSIS OF VARIANCE

$$(y_{ij} - \bar{y}_{..}) = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (\bar{y}_k - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_k + 2\cdot$$

BY ORTHOGONALITY OF EXPERIMENTAL DESIGN

$$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = t\sum_{i=1}^{t} (\bar{y}_{i.} - \bar{y}_{..})^2 + t\sum_{j=1}^{t} (\bar{y}_{.j} - \bar{y}_{..})^2 + t\sum_k (\bar{y}_k - \bar{y})^2 + SSE$$

$$SSTOTAL = SSR + SSC + SST + SSE$$

SSE is obtained by subtraction

For Model w/ Fixed TRT Effects: $E[MST] = \sigma^2 + t\,\theta_t^2$

ANOVA

SOURCE	df	SS	MS	F_0
ROWS	$t-1$	$t\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2$	$MSR = SSR/(t-1)$	$F_0 = \frac{MST}{MSE}$
COLUMNS	$t-1$	$t\sum_j (\bar{y}_{.j} - \bar{y}_{..})^2$	$MSC = SSC/(t-1)$	
TRTS	$t-1$	$t\sum_k (\bar{y}_k - \bar{y}_{..})^2$	$MST = SST/(t-1)$	
ERROR	$(t-1)(t-2)$	SS ERROR	$MSE = SSE/[(t-1)(t-2)]$	
TOTAL	t^2-1	$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2$		

STD Errors for TRT MEANS: $S_{\bar{y}_k} = \sqrt{\frac{MSE}{t}}$

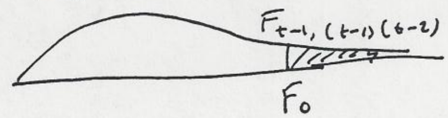
STD Error for Difference between 2 trt means: $S_{\bar{y}_k - \bar{y}_{k'}} = \sqrt{\frac{2MSE}{t}}$

H_0: All treatment means are =

H_A: NOT All TRT MEANS ARE =

T.S. $F_0 = \frac{MST}{MSE}$ RR: $F_0 \geq F_{\alpha,\, t-1,\, (t-1)(t-2)}$

P-VALUE



Relative Efficiency of Column Blocking (uncorrected)

$$S^2_{RCB} = \frac{MS\,Columns + (t-1)MSE}{t} \qquad S^2_{LS} = MSE$$

$$RE_{COL} = \frac{S^2_{RCB}}{S^2_{LS}}$$

RELATIVE EFFICIENCY OF ROW BLOCKING (UNCORRECTED)

$$S^2_{RCB} = \frac{MS\,Rows + (t-1)MSE}{t} \qquad S^2_{LS} = MSE$$

$$RE_{ROW} = \frac{S^2_{RCB}}{S^2_{LS}}$$

When t is small, squares will need to be replicated. Can they be completed in the same rows or columns? See Kuehl pp 287-289 for analysis.

Cochran & Cox (1957) P. 121–123

EXAMPLE – 6 (TRTS!) Wheat Samplers who judge height of wheat shoots

COLUMNS: 6 AREAS TO BE JUDGED (small ≈80 # of shoots

ROWS: ORDER OF JUDGMENT

NOTE: No two judges were in the same area at the same time.

$Y \equiv$ sampler error $\equiv$ difference between mean height of representative shoots judged by sampler and mean height of all shoots in area ($Y > 0 \Rightarrow$ chose too large of shoots)

Treating Samplers as Fixed

AREA (j)

ORDER (i)	1	2	3	4	5	6	$\bar{Y}_{i.}$
I	F +3.5	B +4.2	A +6.7	D +6.6	C +4.1	E +3.8	4.82
II	B +8.9	F +1.9	D +5.8	A +4.5	E +2.4	C +5.8	4.88
III	C +9.6	E +3.7	F −2.7	B +3.7	D +6.0	A +7.0	4.55
IV	D +10.5	C +10.2	B +4.6	E +3.7	A +5.1	F +3.8	6.32
V	E +3.1	A +7.2	C +4.0	F −3.3	B +3.5	D +5.0	3.25
VI	A +5.9	D +7.6	E −0.7	C +3.0	F +4.0	B +8.6	4.73
$\bar{Y}_{.j}$	6.92	5.80	2.95	3.03	4.18	5.67	$\bar{Y}_{..}$ 4.7583
$\bar{Y}_K$	6.07	5.58	6.12	6.92	2.67	1.20	

$\sum\sum Y_{ij}^2 = 1144.73$

$$SS\,TOTAL = \sum_i\sum_j (Y_{ij} - \bar{Y}_{..})^2 = \sum_i\sum_j Y_{ij}^2 - t^2\,\bar{Y}_{..}^2 = \cancel{1144.73}$$

$$1144.73 - 36(4.7583)^2 = 329.63$$

$$SS\,ROWS = 6\left[(4.82 - 4.7583)^2 + \dots + (4.73 - 4.7583)^2\right] = 28.60$$

$$SS\,COLS = 6\left[(6.92 - 4.7583)^2 + \dots + (5.67 - 4.7583)^2\right] = 78.87$$

$$SS\,TRTS = 6\left[(6.07 - 4.7583)^2 + \dots + (1.20 - 4.7583)^2\right] = 155.60$$

$$SS\,ERROR = 329.63 - 28.60 - 78.87 - 155.60 = 66.56$$

8.13

ANOVA

Source	df	SS	MS	F_0	$F^{.05}_{\alpha,\nu_1,\nu_2}$
Rows (Order)	5	28.60	5.72		
Cols (Areas)	5	78.87	15.77		
Trts (Samplers)	5	155.60	31.12	9.35*	2.71
Error	20	66.56	3.33		
Total	35	329.63	—		

* => Conclude not all μ_k are = (Samplers differ)

$$s_{\bar{y}_k - \bar{y}_{k'}} = \sqrt{\frac{2\,MSE}{t}} = \sqrt{\frac{2(3.33)}{6}} = 1.05$$

Bonferroni $\binom{6}{2} = 15$ pairs of samplers: $t_{.025,15,20} = 3.33$

=> Conclude $\mu_k \neq \mu_{k'}$ if $|\bar{y}_k - \bar{y}_{k'}| \geq 3.33(1.05) = 3.50$

Samplers (Bonferroni)

1.20	2.67	5.58	6.07	6.12	6.92
F	E	B	A	C	D

Lines adjoining samplers => Samplers are not significantly different (Note that E & C are 0.05 from being significantly different)

Tukey $q_{.05,6,20} = 4.45$ $\quad HSD = q_{.05,6,20}\sqrt{\frac{MSE}{t}} = 4.45\sqrt{\frac{3.33}{6}} = 3.32$

Samplers (Tukey)

1.20	2.67	5.58	6.07	6.12	6.92
F	E	B	A	C	D

2 Extra pairs are significantly different