p228

M484

62

11) The CRBD model is

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \qquad \text{(additive)}$$

where τ_i is the ith trt effect and $\sum_{i=1}^{k} \tau_i = 0$,

β_j is the jth block effect and $\sum_{j=1}^{b} \beta_j = 0$.

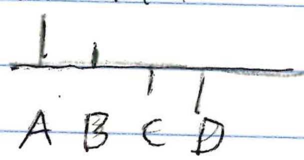
Then $\mu_i = \frac{\mu_{i0}}{b} = \frac{1}{b} \sum_{j=1}^{b} (\mu + \tau_i + \beta_j) = \mu + \tau_i$.

The errors e_{ij} are iid with 0 mean and constant variance σ^2.

12) * Blocking and randomization within blocks can be used to make the iid e_{ij} assumption reasonable.

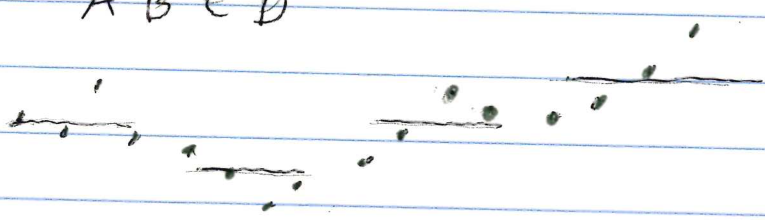
ex) Yields from adjacent identically treated plots tend to be similar, but Yield varies widely and is not iid.

mean a)



"signal" = mean yields

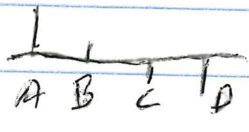
b)



Yields are not iid

blocks B C A D A B D C D C A B D C B A

Filter out noise with solid lines = block means shifted so they add to 0.



estimated "signal" = estimated trt means

Final review 1

ML484

8 or 7 pages 300 pts 15 sheets of notes

see Q10, E1–E3, then Q1–9, old final

1)* Find $\hat{Y}$, CI for β_i, test $H_0\ \beta_i = 0$ E1

2)* MLR Anova F test E1

3)* Partial F test H_0 reduced model is good E1

4)* If you use a table, give the df

df for t table df = n − p

use 1.645 1.96 2.576
90% 95% 99%

if df > 30.

For Partial F test, den df = n−p,

if den df > 60, use df = ∞, otherwise use

den. df closest to n−p.

The num df = $df_R - df_F$ = p−8 = # parameters set to 0 in the reduced model.

See 3 paragraphs below 13) in E1 rev.

5) * Variable selection

Find I_{min} and $C_p(I_{min})+1$; I_I has nomore ← model with fewest predictors such that

predictors than I_{min} and $C_p(I_I) \le C_p(I_{min})+1$.

Models I with fewer predictors than I_I

such that $C_p(I) \le C_p(I_{min})+4$ should also

be examined. See E3 6, 9* E2 5, 8

Give k terms of models I_{min}, I_I including constant

6) log rule $\frac{\max(x_i)}{\min(x_i)} > 10$ → use $\log x_i$

if $\min x_i > 0$

E1 2

7) ladder rule to spread small values, make λ smaller
large larger

E2 9

8) residual plot 

response plot 

MLR

may need to sketch

E1 5

9) one way Anova plots are similar but have
dot plots

10) transformation plots MLR, DOE :

look for plot that is a good response plot.

E2 3 E3 7

434 final review
2

11) * fixed effects one way Anova

find $\hat{\mu}_i = \overline{Y}_{i0}$ and Anova F test E3 /

12) * random effects one way Anova

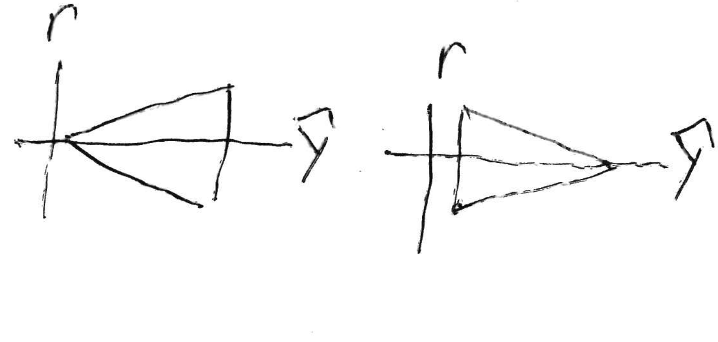
i) because trt levels are a random sample from pop of levels

ii) Anova F test E3

13) Multiple comparisons: groups of <u>pop</u> means that are not significantly different. (μ_1 and μ_2, μ_4 and μ_7 not $\overline{Y}_1$ and $\overline{Y}_2$, $\overline{Y}_4$ and $\overline{Y}_7$)

14) why residual or response plot is bad or find outlier (gap in response plot, plots fill space)

non constant variance



E3 2

15) 2 way ANOVA tests AB interaction

A

B

Q10

2.5

16 * block CRBD a) was blocking useful

$P_{block} < 0.05$ yes

> 0.1 no

between .05 and .1 borderline

b) Anova F test

Q10 2

17)* Given $Y_i = \beta_1 + \beta_2 X_i + e_i$

one of β_1, β_2 known i) find

LS estimator of the unknown

ii) $E(Y_i) = \beta_1 + \beta_2 X_i$ E2 >

see E2 rev 19) and following ex.

There could be other problems.

$\overline{\text{484 final rev}}$ 3

$\cancel{(*)}$ $Y_i = a_i + \beta x_i + e_i$, a_i known

$$Q(\eta) = \sum_{i=1}^{n} (Y_i - a_i - \eta x_i)^2$$

$$E(Y_i | x_i) = a_i + \beta x_i$$

$$\frac{dQ(\eta)}{d\eta} = \sum_{i=1}^{n} 2(Y_i - a_i - \eta x_i)(-x_i) =$$

$$-2 \sum_{i=1}^{n} x_i (Y_i - a_i - \eta x_i)$$

$$= -2 \left[\sum_{i=1}^{n} x_i (Y_i - a_i) - \eta \sum_{i=1}^{n} x_i^2 \right]$$

$$\underbrace{= 2\eta \sum_{i=1}^{n} x_i^2 - 2 \sum_{i=1}^{n} x_i (Y_i - a_i)}_{\frac{dQ(\eta)}{d\eta}} \overset{set}{=} 0 \quad (*)$$

or $\eta \; 2 \sum x_i^2 = 2 \sum x_i (Y_i - a_i)$

$$\text{so } \hat{\eta} = \hat{\beta} = \frac{\sum_{i=1}^{n} x_i (Y_i - a_i)}{\sum_{i=1}^{n} x_i^2}$$

and $\dfrac{d^2 Q(\eta)}{d\eta^2} = 2 \sum_{i=1}^{n} x_i^2 > 0.$

If $x_i \equiv 1$, then $\dfrac{d^2 Q(\eta)}{d\eta^2} = 2n > 0$, $\dfrac{dQ(\eta)}{d\eta} = 2n\eta - 2\sum (Y_i - a_i) \overset{set}{=} 0$

$$\text{so } \hat{\eta} = \hat{\beta} = \frac{\sum (Y_i - a_i)}{n}.$$