Chapter 24

Multifactor Studies - when three or more factors are studied simultaneously

Notation $\mu_{ij\cdot} = \frac{\sum_k \mu_{ijk}}{c}$ $\quad \mu_{i\cdot k} = \frac{\sum_j \mu_{ijk}}{b}$ $\quad \mu_{\cdot jk} = \frac{\sum_i \mu_{ijk}}{a}$

$\mu_{i\cdot\cdot} = \frac{\sum_j \sum_k \mu_{ijk}}{bc}$ $\quad \mu_{\cdot j\cdot} \frac{\sum_i \sum_k \mu_{ijk}}{ac}$ $\quad \mu_{\cdot\cdot k} = \frac{\sum_i \sum_j \mu_{ijk}}{ab}$

$\mu_{\cdot\cdot\cdot} = \frac{\sum_i \sum_j \sum_k \mu_{ijk}}{abc}$

Main effect of the: $i^{th}$ level of factor A: $\alpha_i = \mu_{i\cdot\cdot} - \mu_{\cdot\cdot\cdot}$

$j^{th}$ level of factor B: $\beta_j = \mu_{\cdot j\cdot} - \mu_{\cdot\cdot\cdot}$

$k^{th}$ level of factor C: $\gamma_k = \mu_{\cdot\cdot k} - \mu_{\cdot\cdot\cdot}$

Two factor Interactions
- between factor A at the ith level and factor B at the $j^{th}$ level: $(\alpha\beta)_{ij} = \mu_{ij\cdot} - \mu_{i\cdot\cdot} - \mu_{\cdot j\cdot} + \mu_{\cdot\cdot\cdot}$
- between factor A at the $i^{th}$ level and factor C at the $k^{th}$ level: $(\alpha\gamma)_{ik} = \mu_{i\cdot k} - \mu_{i\cdot\cdot} - \mu_{\cdot\cdot k} + \mu_{\cdot\cdot\cdot}$
- between factor B at the $j^{th}$ level and factor C at the $k^{th}$ level: $(\beta\gamma)_{jk} = \mu_{\cdot jk} - \mu_{\cdot j\cdot} - \mu_{\cdot\cdot k} + \mu_{\cdot\cdot\cdot}$

NOTE: $\sum(\alpha\beta)_{ij} = 0 \quad \forall$ i and j

$\sum(\alpha\gamma)_{ik} = 0 \quad \forall$ k and i

$\sum(\beta\gamma)_{jk} = 0 \quad \forall$ k and j

Three factor Interaction $(\alpha\beta\gamma)_{ijk} = \mu_{ijk} - [\mu_{\cdot\cdot\cdot} + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk}$

NOTE: $\sum_i (\alpha\beta\gamma) = 0 \quad \forall\ j,k$

$\sum_j (\alpha\beta\gamma) = 0 \quad \forall\ i k$

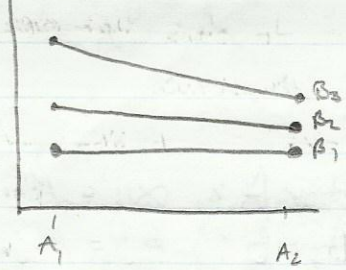
$\sum_k (\alpha\beta\gamma) = 0 \quad \forall\ i j$

**Cell Means** $Y_{ijkm} = \mu_{ijk} + \epsilon_{ijkm}$

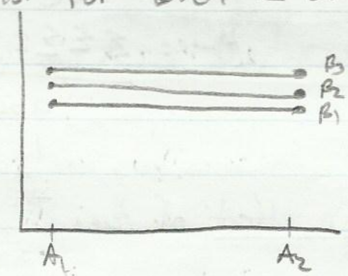
**Factor Effects** $Y_{ijkm} = \mu_{\cdots} + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkm}$

Three Factor

AB Plot for level 1 of C

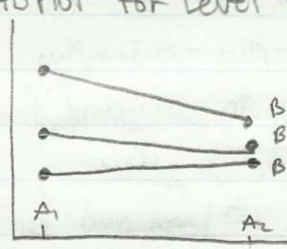



AB Plot for level 2 of C

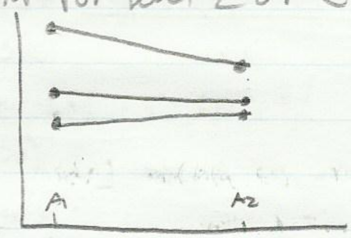



- Here, as the two graphs won't be approximately parallel a three factor interaction will exist.

AB Plot for Level 1 of C




AB Plot for level 2 of C

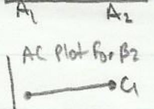



- Here, as the two graphs will be approximately parallel a three factor interaction will not exist

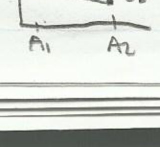
Two Factor

- We can also check these for two factor interactions which we find for AB because of the the curves aren't parallel in each panel
- Now we can check for AC interactions, which we find with a lack of parallel curves, and BC interactions which we don't find as the curves are parallel

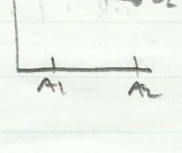
AC Plot for $B_1$



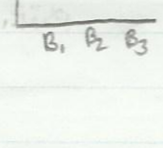

AC Plot for $B_2$



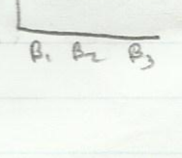

AC Plot for $B_3$




BC plot for $C_1$




BC plot for $C_2$



SAS

```
proc sgscatter noautolegend;  /* Two factor interactions */
    series x= A   y=Y / group=B
    scatter x=A   y=Y / group=B  markerchar=judge
run;
/* You'd have to get separate graphs for each level of C
   a clever data step could do this but sgpanel is easier */
proc sgpanel;  /* three factor interactions
   panelby c / rows=1 columns=k;   /* Note: we can get interaction
   scatter  x= A  y=Y / group=B;       plots by fitting the
   reg      x=A   y=Y / group=B;       appropriate model in
run;                                   glm w/ plots=all; */
proc sgscatter;  /* Main effects Plot */
    plot y*A / loess;
run;
```

Fitting the ANOVA Model

| | |
|---|---|
| $Y_{ijk\cdot} = \sum_m Y_{ijkm}$ | $\bar{Y}_{ijk\cdot} = Y_{ijk\cdot}/n$ |
| $Y_{ij\cdot\cdot} = \sum_k \sum_m Y_{ijkm}$ | $\bar{Y}_{ij\cdot\cdot} = Y_{ij\cdot\cdot}/cn$ |
| $Y_{i\cdot k\cdot} = \sum_j \sum_m Y_{ijkm}$ | $\bar{Y}_{i\cdot k\cdot} = Y_{i\cdot k\cdot}/bn$ |
| $Y_{\cdot jk\cdot} = \sum_i \sum_m Y_{ijkm}$ | $\bar{Y}_{\cdot jk\cdot} = Y_{\cdot jk\cdot}/an$ |
| $Y_{i\cdot\cdot\cdot} = \sum_j \sum_k \sum_m Y_{ijkm}$ | $\bar{Y}_{i\cdot\cdot\cdot} = Y_{i\cdot\cdot\cdot}/bcn$ |
| $Y_{\cdot j\cdot\cdot} = \sum_i \sum_k \sum_m Y_{ijkm}$ | $\bar{Y}_{\cdot j\cdot\cdot} = Y_{\cdot j\cdot\cdot}/acn$ |
| $Y_{\cdot\cdot k\cdot} = \sum_i \sum_j \sum_m Y_{ijkm}$ | $\bar{Y}_{\cdot\cdot k\cdot} = Y_{\cdot\cdot k\cdot}/abn$ |
| $Y_{\cdot\cdot\cdot\cdot} = \sum_i \sum_j \sum_k \sum_m Y_{ijkm}$ | $\bar{Y}_{\cdot\cdot\cdot\cdot} = Y_{\cdot\cdot\cdot\cdot}/abcn$ |

$\hat{\mu}_{ijk} = \bar{Y}_{ijk\cdot} \longrightarrow \hat{Y}_{ijkm} = \bar{Y}_{ijk\cdot}$ w/ $e_{ijkm} = Y_{ijkm} - \hat{Y}_{ijkm} = Y_{ijkm} - \bar{Y}_{ijk\cdot}$

$\hat{\mu}_{\cdot\cdot\cdot} = \bar{Y}_{\cdot\cdot\cdot\cdot}$ $\quad (\alpha\beta)_{ij} = \bar{Y}_{ij\cdot\cdot} + \bar{Y}_{i\cdot\cdot\cdot} + \bar{Y}_{\cdot j\cdot\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$

$\hat{\alpha}_i = \bar{Y}_{i\cdot\cdot\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$ $\quad (\alpha\gamma)_{ik} = \bar{Y}_{i\cdot k\cdot} + \bar{Y}_{i\cdot\cdot\cdot} + \bar{Y}_{\cdot\cdot k\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$

$\hat{\beta}_j = \bar{Y}_{\cdot j\cdot\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$ $\quad (\beta\gamma)_{jk} = \bar{Y}_{\cdot jk\cdot} + \bar{Y}_{\cdot j\cdot\cdot} + \bar{Y}_{\cdot\cdot k\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$

$\hat{\gamma}_k = \bar{Y}_{\cdot\cdot k\cdot} + \bar{Y}_{\cdot\cdot\cdot\cdot}$ $\quad (\alpha\beta\gamma)_{ijk} = \bar{Y}_{ijk\cdot} - \bar{Y}_{ij\cdot\cdot} - \bar{Y}_{i\cdot k\cdot} - \bar{Y}_{\cdot jk\cdot} + \bar{Y}_{i\cdot\cdot\cdot} + \bar{Y}_{\cdot j\cdot\cdot} + \bar{Y}_{\cdot\cdot k\cdot} - \bar{Y}_{\cdot\cdot\cdot\cdot}$

$\boxed{\text{SAS}}$ to fit the ANOVA MODEL

```
proc glm  outstat=full;                 /* Check which interactions
    class  A  B  C;                        we can drop (Type III)*/
    model  Y= A|B|C /solution;
run;
proc glm outstat=reduced;                /* fit the reduced model*/
    class  A  B C;
    model  Y= A  B|C /solution;
run
/* We can test that it's appropriate to use the reduced */
data test1;
    set full reduced;
    if _SOURCE_="ERROR";
proc means data=test1 noprint;
    var ss df;
    output out= test2  min=minss mindf  max=maxss  maxdf;
data nested;
    set test2;
    fstar= ((maxss-minss) /(maxdf - mindf))/(minss/mindf);
    pvalue= 1 - cdf('f', fstar, maxdf-mindf, mindf);
proc print;
run;          /* Look at  P-value */
```

- Sum of Squares and Mean squared errors are outlined in table 24.5 on pg. 1006
- Check Residuals the same way: * Residuals vs each treatment
  * Normal Probability Plot
  * Residuals vs predicted val

## More sums of squares

SSTO = SSTreatment + SSE

$$\equiv \sum_i\sum_j\sum_k\sum_m (Y_{ijkm} - \bar{Y}_{....})^2 = n\sum_i\sum_j\sum_k (\bar{Y}_{ijk.} - \bar{Y}_{....}) + \sum_i\sum_j\sum_k\sum_m e^2_{ijkm}$$

SSTreatment = SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC

| | | degrees of freedom |
|---|---|---|
| where | $SSA = nbc\sum_i(\bar{Y}_{i...} - \bar{Y}_{....})^2$ | $a-1$ |
| | $SSB = nac\sum_j(\bar{Y}_{.j..} - \bar{Y}_{....})^2$ | $b-1$ |
| | $SSC = nab\sum_k(\bar{Y}_{..k.} - \bar{Y}_{....})^2$ | $c-1$ |
| | $SSAB = nc\sum_i\sum_j(\bar{Y}_{ij..} - \bar{Y}_{i...} - \bar{Y}_{.j..} + \bar{Y}_{....})^2$ | $(a-1)(b-1)$ |
| | $SSAC = nb\sum_i\sum_k(\bar{Y}_{i.k.} - \bar{Y}_{i...} - \bar{Y}_{..k.} + \bar{Y}_{....})^2$ | $(a-1)(c-1)$ |
| | $SSBC = na\sum_j\sum_k(\bar{Y}_{.jk.} - \bar{Y}_{.j..} - \bar{Y}_{..k.} + \bar{Y}_{....})^2$ | $(b-1)(c-1)$ |
| | $SSABC = n\sum_i\sum_j\sum_k(\bar{Y}_{ijk.} - \bar{Y}_{ij..} - \bar{Y}_{i.k.} - \bar{Y}_{.jk.} + \bar{Y}_{i...} + \bar{Y}_{.j..} + \bar{Y}_{..k.} - \bar{Y}_{....})^2$ | $(a-1)(b-1)(c-1)$ |

SSTO = SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC + SSE

## F-Test for factor effect

$H_0$: All of this factor level = 0  ie: $(\alpha\beta\gamma)_{ijk} = 0$

$H_a$: Not $H_0$

$F^* = \dfrac{\text{Associated Mean Square}}{\text{Mean Square Error}}$  ie: $F^* = \dfrac{MSABC}{MSE}$

Reject if $F^* > F[1-\alpha, \text{degrees of freedom}, (n-1)abc]$ ie Reject if $F^* > F[1-\alpha, (a-1)(b-1)(c-1), (n-1)abc]$

## Strategy for analysis

1.) Do three factor interactions exist?
2.) If not do two factor interactions exist?
3.) If not can we drop single main effect?
4.) If yes ↳ 1, can we make a transformation to make it unimportant
5.) If not analyze three factors jointly in terms of $\mu_{ijk}$
6.) If theres just one important two factor analyze the effects jointly in terms of $\mu_{ij.}$, $\mu_{.jk}$, $\mu_{i.k}$. Analyze third effect separately

7.) If there are two or three important two-factor interactions analyze the three factors jointly in terms of $\mu_{ijk}$.

** There are exceptions to this strategies **

Estimation of factor level Mean $\hat{\mu}_{i..} = \bar{Y}_{i...}$

w/ variance $s^2(\bar{Y}_{i...}) = \frac{MSE}{nbc}$

Confidence intervals for factor level mean: $\bar{Y}_{i...} \pm t[1-\alpha/2, (n-1)abc]\, s(\bar{Y}_{i...})$

Contrast of factor level means: $L = \Sigma c_i \mu_{i..}$ where $\Sigma c_i = 0$

Confidence intervals for contrasts: $\hat{L} \pm t[1-\alpha/2; (n-1)abc]\, s(\hat{L})$

w/ $\hat{L} = \Sigma c_i \bar{Y}_{i...}$

$s^2(\hat{L}) = \frac{MSE}{nbc} \Sigma c_i^2$

Hypothesis Test for contrasts $H_0: L = 0$

$H_a: L \neq 0$

$t^* = \frac{\hat{L}}{s(\hat{L})}$ ; Reject when $> t[1-\alpha/2; (n-1)(abc)]$

Multiple Contrasts Conf. Interval

$\hat{L} \pm \frac{1}{\sqrt{2}} q[1-\alpha; a, (n-1)abc]$ TUKEY

$\hat{L} \pm (a-1) F[1-\alpha, a-1, (n-1)(abc)]$ Scheffe

$\hat{L} \pm t[1-\alpha/2g; (n-1)abc]$ BONFERRONI

Multiple Contrasts Hypothesis Test

$q^* = \frac{\sqrt{2}\,\hat{D}}{s(\hat{D})}$ reject if $> q[1-\alpha, a, (n-1)abc]$

$F^* = \frac{\hat{L}^2}{(a-1)s^2(\hat{L})}$ reject if $> F[1-\alpha, a-1, (n-1)abc]$

$t^* = \frac{\hat{L}}{s(\hat{L})}$ reject if $> t[1-\alpha/2g, (n-1)abc]$

$\boxed{\text{SAS}}$

```
proc glm;  /*All pairwise differences*/
   class A B C;
   model Y = A B|C;
   lsmeans A / pdiff adjust= tukey alpha=.05 cl;
   lsmeans B*C/pdiff adjust= tukey
run;
```

```
proc glimmix; /* Contrasts */
   class A B C;
   model Y= A B C B*C;
                                   + scheffe
   lsmestimate A "A" 1 -1 / adjust= bonferroni alpha=.05 cl;
   lsmestimate B*C "d1" -1 1 0 0,
                                          + scheffe
                   "d2" 0 0 -1 1 / adjust= bonferroni alpha=.05 cl;
run;
```

<u>Skip 24.6 & 24.7</u>