Chapter 19 Two Factor studies

Factor Level Means: $\mu_{\cdot j} = \sum_{i=1}^{a} \mu_{ij}/a$, $\mu_{i\cdot} = \sum_{j=1}^{b} \mu_{ij}/b$

overall mean: $\mu_{\cdot\cdot} = \sum_i \sum_j \mu_{ij}/ab = \sum_i \mu_{i\cdot}/a = \sum_j \mu_{\cdot j}/b$

Main Effects: The different between each factor level mean and the overall mean

$\beta_1 = \mu_{\cdot 1} - \mu_{\cdot\cdot}$ $\qquad \alpha_1 = \mu_{1\cdot} - \mu_{\cdot\cdot}$

thus $\beta_j = \mu_{\cdot j} - \mu_{\cdot\cdot}$ $\qquad \alpha_i = \mu_{i\cdot} - \mu_{\cdot\cdot}$

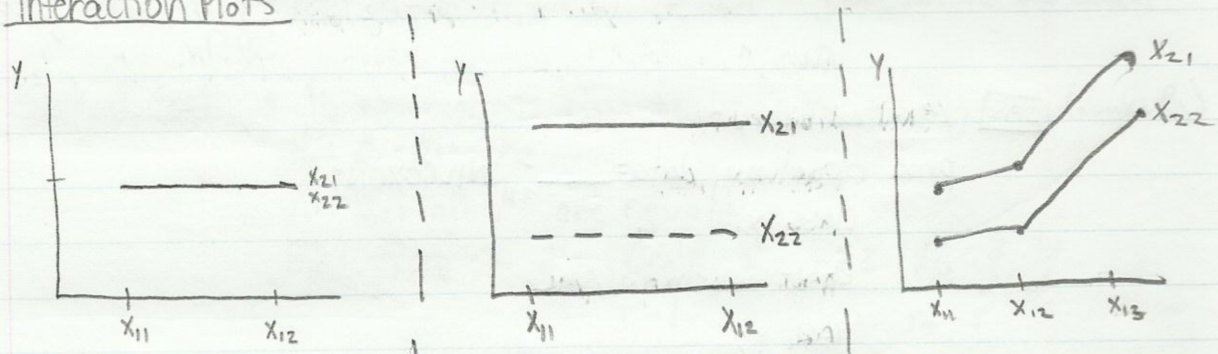
Note $\sum_{j=1}^{b} \beta_j = 0$ $\qquad \sum_{i=1}^{a} \alpha_i = 0$

Additive factor Effects: $\mu_{ij} = \mu_{\cdot\cdot} + \alpha_i + \beta_j = \mu_{i\cdot} + \mu_{\cdot j} - \mu_{\cdot\cdot}$

↑ If all treatment means can be written like this we say that the factors do not interact

↑ We can also see this in an interaction plot

Interaction Plots



* Both factor levels don't affect y
* $X_{1\cdot}$ has no effect on y but $X_{2\cdot}$ does
* Both $X_{1\cdot}$ and $X_{2\cdot}$ affect y

** Interactions exist if drastic deviation from parallel lines exist

SAS
```
proc glm data=___ plots=all;
    class x1 x2;
    model y=x1 x2/solution;
run;
```

- If interaction plots suggest an interaction exists we can try transforming

① $\log \mu_{ij} = \log \mu_{..} + \log(\alpha_i) + \log(\beta_j)$ for $\mu_{ij} = \mu_{..}\,\alpha_i \beta_j$

② $\sqrt{\mu_{ij}} = \sqrt{\alpha_i} + \sqrt{\beta_j}$ for $\mu_{ij} = (\sqrt{\alpha_i} + \sqrt{\beta_j})^2$

$\boxed{\text{SAS}}$ Same as 704, a data step and refit with new variables

- Anova Model I: - All treatment sample sizes are equal
  - All treatment means are of equal importance

  Factor A → a levels Factor B → b levels

  $n_T = abn$

  CELL MEANS: $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$

  $E(Y_{ijk}) = \mu_{ij}$ $\sigma^2(Y_{ijk}) = \sigma^2(\epsilon_{ijk}) = \sigma^2$

  $Y_{ijk} \sim N(\mu_{ij}, \sigma^2)$

  Factor Effects $(\alpha\beta)_{ij} = \mu_{ij} - (\mu_{..} + \alpha_i + \beta_j)$

  $\equiv \mu_{ij} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij}$

  where $\mu_{..} = \sum_i \sum_j \mu_{ij}/ab$, $\alpha_i = \mu_{i.} - \mu_{..}$, $\beta_j = \mu_{.j} - \mu_{..}$, $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$

$\boxed{\text{SAS}}$

```
proc glm;
class x1 x2;
model y= x1 x2 x1*x2;
means x1 x2 x1*x2;
run;
```

Output: $\hat{\mu}_{..} = \bar{Y}_{...}$

$\hat{\alpha}_i = \bar{Y}_{i..} - \bar{Y}_{...}$

$\hat{\beta}_j = \bar{Y}_{.j.} - \bar{Y}_{...}$

$\widehat{(\alpha\beta)}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$

$e_{ijk} = Y_{ijk} - \bar{Y}_{ij.}$ → Check residual & normal prob plots

$SSTO = SSTR + SSE = \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{...})^2$ $df = nab - 1$

$SSTR = n \sum_i \sum_j (\bar{Y}_{ij.} - \bar{Y}_{...})^2$ $df = ab - 1$

$SSE = \sum_i \sum_j \sum_k e_{ijk}^2$ $df = (n-1)ab$

$MSTr = \dfrac{SSTr}{ab-1}$

$MSE = \dfrac{SSE}{ab(n-1)}$

<u>Test for Interactions</u> H_0: all $(\alpha\beta)_{ij} = 0$

H_a: not H_0

$F^* = \dfrac{MSAB}{MSE}$

reject $F^* > F[1-\alpha; a-1, (n-1)ab]$

$SSA = nb\,\Sigma(\bar{Y}_{i..} - \bar{Y}_{...})^2$

$SSB = na\,\Sigma(\bar{Y}_{.j.} - \bar{Y}_{...})^2$

$SSAB = n\,\Sigma\Sigma(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$

$MSA = SSA/(a-1)$

$MSB = SSB/(b-1)$

$MSAB = SSAB/[(a-1)(b-1)]$

* Look at F tests of interactions in type III table

<u>Test for Factor Main Effects</u>

$H_0: \mu_{1.} = \mu_{2.} = \dots \mu_{a.}$ | $H_0: \mu_{.1} = \mu_{.2} = \dots \mu_{.b}$

$\equiv H_0: \alpha_1 = \alpha_2 = \dots = \alpha_a = 0$ | $H_0: \beta_1 = \beta_2 = \dots = \beta_b = 0$

H_a: not H_0 | H_a: not H_0

$F^* = \dfrac{MSA}{MSE}$ | $F^* = \dfrac{MSB}{MSE}$

reject if $F^* > F[1-\alpha, a-1, (n-1)ab]$ | $F^* > F[1-\beta, b-1, (n-1)ab]$

* Look at F tests of factor levels in type III table

<u>Confidence Intervals</u> for $\mu_{i.}$: $\bar{Y}_{i..} \pm t[1-\alpha/2, (n-1)ab]\, s(\bar{Y}_{i..})$

for $\mu_{.j}$: $\bar{Y}_{.j.} \pm t[1-\alpha/2, (n-1)ab]\, s(\bar{Y}_{.j.})$

<u>Unbiased point estimators</u>: $\hat{\mu}_{i..} = \bar{Y}_{i..}$

$\hat{\mu}_{.j} = \bar{Y}_{.j.}$

$\sigma^2(\bar{Y}_{i..}) = \sigma^2/bn$ | $s^2(\bar{Y}_{i..}) = MSE/bn$

$\sigma^2(\bar{Y}_{.j.}) = \sigma^2/an$ | $s^2(\bar{Y}_{.j.}) = MSE/an$

<u>Contrasts</u> $L = \Sigma c_i \mu_i$ where $\Sigma c_i = 0$ | $L = \Sigma c_j \mu_{.j}$ where $\Sigma c_j = 0$

$\hat{L} = \Sigma c_i \bar{Y}_{i..}$ | $\hat{L} = \Sigma c_j \bar{Y}_{.j}$

$\sigma^2(\hat{L}) = \Sigma c_i^2 \sigma^2(\bar{Y}_{i..}) = \dfrac{\sigma^2}{bn} \Sigma c_i^2$ | $\sigma^2(\hat{L}) = \dfrac{\sigma^2}{an} \Sigma c_j^2$

$s^2(\hat{L}) = \dfrac{MSE}{bn} \Sigma c_i^2$ | $s^2(\hat{L}) = \dfrac{MSE}{an} \Sigma c_j^2$

$\hat{L} \pm t[1-\alpha/2, (n-1)ab]$

Pairwise Comparisons

* **Tukey** - for all pairwise comparisons

$$D = \mu_{i\cdot\cdot} - \mu_{i'\cdot\cdot}$$
$$\hat{D} = \bar{Y}_{i\cdot\cdot} - \bar{Y}_{i'\cdot\cdot}$$
$$s^2(\hat{D}) = {}^{2MSE}/_{bn}$$
$$T = \frac{1}{\sqrt{2}} q[1-\alpha, a, (n-1)ab]$$

CI: $\hat{D} \pm Ts(\hat{D})$

Hyp test: $H_0: D = \mu_{i\cdot\cdot} - \mu_{i'\cdot\cdot} = 0$

$H_a: D = \mu_{i\cdot\cdot} - \mu_{i'\cdot\cdot} \neq 0$

$q^* = \sqrt{2}\hat{D}/s(\hat{D})$

Reject if $|q^*| > q[1-\alpha, a, (n-1)ab]$

* **Bonferroni** - for only a few pairwise comparisons or small contrasts

$B = t[1-\alpha/2g; (n-1)ab]$ $\ni$ g = number of statements in family

$t^* = \hat{D}/s(\hat{D})$ $\quad t^* = \hat{L}/s(\hat{L})$

Reject if $|t^*| > t[1-\alpha/2g; (n-1)ab]$

* **Scheffé** for a large # of contrasts

$\hat{L} \pm S\, s(\hat{L})$

where $S^2 = (a-1)F[1-\alpha, a-1, (n-1)ab]$

$H_0: L = 0$

$H_a: L \neq 0$

for $\mu_{i\cdot\cdot}$ $\begin{cases} F^* = \dfrac{\hat{L}^2}{(a-1)s^2(\hat{L})} \\ \text{Reject if } F^* > F[1-\alpha, a-1, (n-1)ab] \end{cases}$

for $\mu_{\cdot j}$ $\begin{cases} F^* = \dfrac{\hat{L}^2}{(b-1)s^2(\hat{L})} \\ \text{Reject if } F^* > F[1-\alpha, b-1, (n-1)ab] \end{cases}$

$\boxed{\text{SAS}}$

- proc glm; /* all pairwise */
  class x_1 x_2;
  model $y = x_1\ x_2\ x_1{*}x_2$;
  lsmeans x_1 / pdiff adjust=tukey alpha=.01 cl;
  run;
- proc glimmix; /* linear combos and contrasts
  class x_1 x_2;
  model $y = x_1\ x_2\ x_1{*}x_2$;
  lsmestimate x_1 "$\mu_1 - \mu_2$" 1 -1 0 / adjust= bonferroni / scheffe / t alpha=.01 cl;
  run;