ANOVA Notes

3a

Assessment of uncertainty from text

replicate for exp $k = 1, 2, 3$ ($\uparrow n$)

Y_{ijk}

Factor A
$i = 1 = \text{low}$
$i = 2 = \text{high} = a$

Factor B
$j = 1 = \text{low}$
$j = 2 = \text{high} = b$

Notation

Factor A ↓ / Factor B →		1 Low	b High
1	Low	y_{111} y_{112} y_{113} 28, 25, 27	y_{121} y_{122} y_{123} 18, 19, 23
a	High	y_{211} y_{212} y_{213} 36, 32, 32	$y_{221}, y_{222}, y_{223}$ 31, 30, 29

dont confuse
Main Effect & F calculations

a = sum of

a = # levels of A

$a = 2$
$b = 2$
$n = 3$

$i=1$ A=low
$i=2$ A=high

$$\bar{Y}_{i\cdot\cdot} = \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} = \text{row average} = \frac{1}{b}\sum_{j=1}^{b}\left[\frac{1}{n}\sum_{k=1}^{n} y_{ijk}\right]$$

$$\bar{Y}_{\cdot j\cdot} = \frac{1}{an}\sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk} = \text{column average}$$

5th ed page 128

eq (5-3)

$$\bar{Y}_{ij\cdot} = \frac{1}{n}\sum_{k=1}^{n} y_{ijk} = \text{cell average}$$

$$y_{\cdot\cdot\cdot} = \frac{1}{abn}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} = \text{grand average}$$

pg 160, Table 5-3

Eq (5-5)

And

Mean squares

Sum squares

$$MS_A = \frac{SS_A}{a-1} \Rightarrow SS_A = bn\sum_{i=1}^{a}(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2 \equiv$$ Sum square residuals row ave wrt grand average

$$MS_B = \frac{SS_B}{b-1} \Rightarrow SS_B = an\sum_{j=1}^{b}(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2 \equiv$$ sum square col residual wrt grand avg

$$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)} \Rightarrow SS_{AB} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot\cdot})^2 \quad ??$$

$$MS_E = \frac{SS_E}{ab(n-1)} \Rightarrow SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{Y}_{ij\cdot})^2 \equiv$$ Sum square cell residual wrt cell ave

For our example

$$\bar{Y}_{1..} = \frac{1}{2}\left[\underbrace{\frac{1}{3}(28+25+27)}_{j=1} + \underbrace{\frac{1}{3}(18+19+23)}_{j=2}\right] = 23.34$$

row 2

$$\bar{Y}_{2..} = \frac{1}{2}\left[\frac{1}{3}(36+32+32) + \frac{1}{3}(31+30+29)\right] = 31.66$$

row averages

col avgs

$$\bar{Y}_{.1.} \xleftarrow{\text{col 1}} = \frac{1}{2}\left[\underbrace{\frac{1}{3}(28+25+27)}_{i=1} + \frac{1}{3}(36+32+32)\right] = 30.0$$

$$\bar{Y}_{.2.} \xleftarrow{\text{col 2}} = \frac{1}{2}\left[\frac{1}{3}(18+19+23) + \frac{1}{3}(31+30+29)\right] = 25.0$$

col averages

$$\bar{Y}_{11.} = \frac{1}{3}\left[(28+25+27)\right] = 26.67 \qquad \bar{Y}_{12.} = \frac{1}{3}\left[18+19+23\right] = 20.0$$

$$\bar{Y}_{21.} = \frac{1}{3}(36+32+32) = 33.3 \qquad \bar{Y}_{22.} = \frac{1}{3}\left[31+30+29\right] = 30.0$$

Cell averages

$$\bar{Y}_{...} = \frac{1}{2\cdot 2\cdot 3}\left[28+25+27+18+19+23+36+32+32+31+30+29\right] = 27.5$$

grand average

Mean Squares

DOF = 1 $\quad MS_A = \frac{2\cdot 3}{2-1}\left[(\bar{Y}_{1..}-\bar{Y}_{...})^2 + (\bar{Y}_{2..}-\bar{Y}_{...})^2\right] = \frac{6}{2-1}\left[(23.34-27.5)^2 + (31.66-27.5)^2\right] = 207.67$

DOF = 1 $\quad MS_B = \frac{2\cdot 3}{2-1}\left[(\bar{Y}_{.1.}-\bar{Y}_{...})^2 + (\bar{Y}_{.2.}-\bar{Y}_{...})^2\right] = \frac{6}{2-1}\left[(30-27.5)^2 + (25.0-27.5)^2\right] = 75.0$

DOF = 1 $\quad MS_{AB} = \frac{3}{(2-1)(2-1)}\Big[(\bar{Y}_{11.}-\bar{Y}_{1..}-\bar{Y}_{.1.}+\bar{Y}_{...})^2 + (\bar{Y}_{12.}-\bar{Y}_{1..}-\bar{Y}_{.2.}+\bar{Y}_{...})^2$
$\quad + (\bar{Y}_{21.}-\bar{Y}_{2..}-\bar{Y}_{.1.}+\bar{Y}_{...})^2 + (\bar{Y}_{22.}-\bar{Y}_{2..}-\bar{Y}_{.2.}+\bar{Y}_{...})^2\Big]$

$$= \frac{3}{1\cdot 1}\Big[(26.67-23.34-30+27.5)^2 + (20-23.34-25.0+27.5)^2 + (33.3-31.66-30.0+27.5)^2 + (30.0-31.66-25.0+27.5)^2\Big]$$

$$= 3\left[0.68+0.706+0.740+0.706\right] = 2.11$$

DOF = 8 $\quad MS_E = \frac{1}{2\cdot 2(3-1)}\Big[(28-26.67)^2+(25-26.67)^2+(27-26.67)^2+(18-20.0)^2+(19-20.0)^2+(23-20.0)^2$
$\quad +(36-33.3)^2+(32-33.3)^2+(32-33.3)^2+(31-30.0)^2+(30-30.0)^2+(29-30.0)^2\Big]$

$$MS_E = 3.92$$

Table 5-3 The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a - 1$	$MS_A = \dfrac{SS_A}{a-1}$	$F_0 = \dfrac{MS_A}{MS_E}$
B treatments	SS_B	$b - 1$	$MS_B = \dfrac{SS_B}{b-1}$	$F_0 = \dfrac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \dfrac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \dfrac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \dfrac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn - 1$		

Don't use

a = # levels of A

b = # levels of B

Row, Col and Cell and Grand Averages

$y_{1..} := \frac{1}{6}\cdot(28+25+27+18+19+23) \quad y_{1..} = 23.333$

$y_{2..} := \frac{1}{6}\cdot(36+32+32+31+30+29) \quad y_{2..} = 31.667$

$y_{.1.} := \frac{1}{6}\cdot(28+25+27+36+32+32) \quad y_{.1.} = 30$

$y_{.2.} := \frac{1}{6}\cdot(18+19+23+31+30+29) \quad y_{.2.} = 25$

$y_{11.} := \frac{1}{3}\cdot(28+25+27) \quad y_{11.} = 26.667 \qquad y_{21.} := \frac{1}{3}\cdot(36+32+32) \quad y_{21.} = 33.333$

$y_{12.} := \frac{1}{3}\cdot(18+19+23) \quad y_{12.} = 20 \qquad y_{22.} := \frac{1}{3}\cdot(31+30+29) \quad y_{22.} = 30$

$y_{...} := \frac{1}{12}\cdot(28+25+27+18+19+23+36+32+32+31+30+29) \quad y_{...} = 27.5$

Row Col Interaction Sum Square Residuals

$$SS_A = 2\bullet 3\left[(\bar{y}_{1\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2+(\bar{y}_{2\bullet\bullet}-\bar{y}_{\bullet\bullet\bullet})^2\right]$$
$$SS_B = 2\bullet 3\left[(\bar{y}_{\bullet1\bullet}-\bar{y}_{\bullet\bullet\bullet})^2+(\bar{y}_{\bullet2\bullet}-\bar{y}_{\bullet\bullet\bullet})^2\right]$$
$$SS_{AB} = 3\bullet\left[\begin{array}{c}(\bar{y}_{11\bullet}-\bar{y}_{1\bullet\bullet}-\bar{y}_{\bullet1\bullet}+\bar{y}_{\bullet\bullet\bullet})^2+(\bar{y}_{12\bullet}-\bar{y}_{1\bullet\bullet}-\bar{y}_{\bullet2\bullet}+\bar{y}_{\bullet\bullet\bullet})^2+(\bar{y}_{21\bullet}-\bar{y}_{2\bullet\bullet}-\bar{y}_{\bullet1\bullet}+\bar{y}_{\bullet\bullet\bullet})^2\\ +(\bar{y}_{22\bullet}-\bar{y}_{2\bullet\bullet}-\bar{y}_{\bullet2\bullet}+\bar{y}_{\bullet\bullet\bullet})^2\end{array}\right]$$
$$SS_E = \left[\begin{array}{c}(y_{111}-y_{11\bullet})^2+(y_{112}-y_{11\bullet})^2+(y_{113}-y_{11\bullet})^2+(y_{121}-y_{12\bullet})^2+(y_{122}-y_{12\bullet})^2+(y_{123}-y_{12\bullet})^2+\\ (y_{211}-y_{21\bullet})^2+(y_{212}-y_{21\bullet})^2+(y_{213}-y_{21\bullet})^2+(y_{221}-y_{22\bullet})^2+(y_{222}-y_{22\bullet})^2+(y_{223}-y_{22\bullet})^2\end{array}\right]$$

$SS_A := 2\cdot 3\cdot\left[(23.33-27.5)^2+(31.667-27.5)^2\right] \quad SS_A = 208.517$

$SS_B := 2\cdot 3\cdot\left[(30-27.5)^2+(25-27.5)^2\right] \quad SS_B = 75$

25

coded right?

$SS_{AB} := 3\cdot\left[(26.67-23.33-30+27.5)^2+(20-23.33-31.67+27.5)^2+(33.33-31.67-30+27.5)^2 \ldots \right.$
$\left. +(30-31.67-25+27.5)^2\right]$

$SS_{AB} = \cancel{175.05}$

$SS_E :=$ ▪

$$SS_E := \begin{bmatrix} (28-26.67)^2+(25-26.67)^2+(27-26.67)^2+(18-20)^2+(19-20)^2+(23-20)^2 \ldots \\ +(36-33.33)^2+(32-33.33)^2+(32-33.33)^2+(31-30)^2+(30-29)^2+(29-29)^2 \end{bmatrix}$$

$$SS_E = 31.333$$

$$MS_A := \frac{SS_A}{2-1} \qquad MS_A = 208.517 \qquad MS_B := \frac{SS_B}{2-1} \qquad MS_B = 75$$

$$MS_{AB} := \frac{SS_{AB}}{(2-1)\cdot(2-1)} \qquad MS_{AB} = \cancel{175.05} \qquad MS_E := \frac{SS_E}{2\cdot 2\cdot(3-1)} \qquad MS_E = 3.917$$

$$FA := \frac{MS_A}{MS_E} \qquad FA = 53.238 \qquad DOF\,A = a-1 = 1$$

$$DOF\,B = b-1 = 1$$ see table 5.3 pg 180

$$FB := \frac{MS_B}{MS_E} \qquad FB = 19.149 \qquad DOF\,AB = (a-1)(b-1) = 1$$

$$FAB := \frac{MS_{AB}}{MS_E} \qquad FAB = 44.694 \qquad DOF\,E = ab(n-1) = 2\cdot 2(3-1) = 8$$

$$Pr\{FA > F_{0.05,1,8}\} = 5\% \quad F_{0.05,1,8} = 5.32; \quad Pr\{FA > 5.32\} = 5\%$$

$$Pr\{FA > F_{0.01,1,8}\} = 1\%; \quad F_{0.01,1,8} = 11.26; \quad Pr\{FA > 11.26\} = 1\%$$

There are two hypothesis

(1) - Difference in variance is caused by factor A

(2) - Difference in variance is a statistical fluke

because

$\therefore$ Since the probability that a statistical fluke has occurred is less than 1%, we choose the alternate that the effect of factor A is statistically significant.