ORIGIN ≡ 1

Nested Two-way ANOVA Model - Balanced Case

W. Stein

Nested ANOVA designs occur when, either by efficiency of design or by necessity, not all combinations of factors are studied. Nested design in a two-way ANOVA (involving two factors) permits analysis of the independent factor (factor A below) in the same way as a crossed two-way ANOVA. The difference lies in factor B nested within A. Here effects of different cell blocks within single treatment levels of A are assessed much as if in single factor ANOVA. Example below comes from Chapter 26 in Kuter et al. (KNNL) *Applied Linear Statistical Models* 5th Edition. See particularly Table 26.2 for discussion of how this design relates to one-way (single-factor) ANOVA.

Data Structure:

Data are structured such that different levels of Nested Factor B occur only within single levels of independent Factor A. Other possible combinations are not studied.

Let index i,j indicate the ith row (treatment classes of Variable A) and jth column (treatment classes of Variable B)

Treatment Classes of Variable A:	Nested Two-Way ANOVA					
	Treatment Classes of Variable B Nested Within A:					
	#1	#2	#3	#4	#5	#6
#1	n	n				
#2			n	n		
#3					n	n
Each cell consists of n replicates with means $Ybar_{ij}$						

Also let: **$Ybar_{i.}$ = mean over all columns for A_i.**
$Ybar_{ij}$ = mean within data blocks.
$Ybar_{..}$ = overall mean.

Model:

Note: this model has no associated test for interactions because of the nested design.

$$Y_{i,jk} = \mu.. + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

where:

$\mu..$ is a constant = grand mean of all objects.
α_i are effect coefficients for classes i in Variable A.
β_j are nested coefficients for classes j of Variable B in A.
ε_{ijk} is the error term specific to each object i,j,k
i = 1 to a, j = 1 to b, k = 1 to n, N = abn

Restrictions:

$$\sum_i \alpha_i := 0 \qquad \sum_j \beta_{j\_in\_i} := 0$$

Assumptions:

– ε_{ijk} are a random sample ~ N(0,σ^2)
– variance is homogeneous across cells

KNNL dataset Table 26.1

Example:

K := READPRN("c:/2008LinearModelsData/TrainingSchoolR.txt")

$$K = \begin{pmatrix} 1 & 25 & 1 & 1 \\ 2 & 29 & 1 & 1 \\ 3 & 14 & 1 & 2 \\ 4 & 11 & 1 & 2 \\ 5 & 11 & 2 & 1 \\ 6 & 6 & 2 & 1 \\ 7 & 22 & 2 & 2 \\ 8 & 18 & 2 & 2 \\ 9 & 17 & 3 & 1 \\ 10 & 20 & 3 & 1 \\ 11 & 5 & 3 & 2 \\ 12 & 2 & 3 & 2 \end{pmatrix}$$

Variable Assignment:

$Y := K^{\langle 2 \rangle}$ **< response (dependent) variable Y**

	B_1	B_2
A_1	$Y_{11} := \begin{pmatrix} 25 \\ 29 \end{pmatrix}$	$Y_{12} := \begin{pmatrix} 14 \\ 11 \end{pmatrix}$
A_2	$Y_{21} := \begin{pmatrix} 11 \\ 6 \end{pmatrix}$	$Y_{22} := \begin{pmatrix} 22 \\ 18 \end{pmatrix}$
A_3	$Y_{31} := \begin{pmatrix} 17 \\ 20 \end{pmatrix}$	$Y_{32} := \begin{pmatrix} 5 \\ 2 \end{pmatrix}$

< response blocks

$$Y_{blocks} := \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \\ Y_{31} & Y_{32} \end{pmatrix}$$

a := 3 b := 2 n := 2 N := length(Y)

Means:

$\mu_{dotdot} := \text{mean}(Y)$ $\qquad \mu_{dotdot} = 15$ **< grand mean**

$$\mu := \begin{pmatrix} \text{mean}(Y_{11}) & \text{mean}(Y_{12}) \\ \text{mean}(Y_{21}) & \text{mean}(Y_{22}) \\ \text{mean}(Y_{31}) & \text{mean}(Y_{32}) \end{pmatrix} \qquad \mu = \begin{pmatrix} 27 & 12.5 \\ 8.5 & 20 \\ 18.5 & 3.5 \end{pmatrix}$$

< block means

$$\mu_{idot} := \begin{pmatrix} \text{mean}(\mu_{1,1}, \mu_{1,2}) \\ \text{mean}(\mu_{2,1}, \mu_{2,2}) \\ \text{mean}(\mu_{3,1}, \mu_{3,2}) \end{pmatrix} \qquad \mu_{idot} = \begin{pmatrix} 19.75 \\ 14.25 \\ 11 \end{pmatrix}$$

< Independent Factor A means (row means)

Model Parameters:

$\alpha := \mu_{idot} - \mu_{dotdot}$

$$\alpha = \begin{pmatrix} 4.75 \\ -0.75 \\ -4 \end{pmatrix}$$

< α coefficients for Independent Factor A

$$\beta_{BinA} := \begin{pmatrix} \mu_{1,1} - \mu_{idot_1} & \mu_{1,2} - \mu_{idot_1} \\ \mu_{2,1} - \mu_{idot_2} & \mu_{2,2} - \mu_{idot_2} \\ \mu_{3,1} - \mu_{idot_3} & \mu_{3,2} - \mu_{idot_3} \end{pmatrix} \qquad \beta_{BinA} = \begin{pmatrix} 7.25 & -7.25 \\ -5.75 & 5.75 \\ 7.5 & -7.5 \end{pmatrix}$$

< β coefficients for Nested Factor B in A

Fitted Values & Residuals:

$$Y = \begin{pmatrix} 25 \\ 29 \\ 14 \\ 11 \\ 11 \\ 6 \\ 22 \\ 18 \\ 17 \\ 20 \\ 5 \\ 2 \end{pmatrix} \qquad \mu_{vec} := \begin{pmatrix} 27 \\ 27 \\ 12.5 \\ 12.5 \\ 8.5 \\ 8.5 \\ 20 \\ 20 \\ 18.5 \\ 18.5 \\ 3.5 \\ 3.5 \end{pmatrix} \qquad e := Y - \mu_{vec} \qquad e = \begin{pmatrix} -2 \\ 2 \\ 1.5 \\ -1.5 \\ 2.5 \\ -2.5 \\ 2 \\ -2 \\ -1.5 \\ 1.5 \\ 1.5 \\ -1.5 \end{pmatrix}$$

< e = residuals

^ means for each block in vector form associated with each value of Y

Sums of Squares:

Σ form:

$$SSA := b \cdot n \cdot \sum_{i=1}^{a} (\alpha_i)^2 \qquad SSA = 156.5$$

$$SSB_{inA} := n \cdot \sum_{i=1}^{a} \sum_{j=1}^{b} \left(\beta_{BinA_{i,j}}\right)^2 \qquad SSB_{inA} = 567.5$$

$$SSE := \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \left[\left(Y_{blocks_{i,j}}\right)_k - \mu_{i,j}\right]^2 \qquad SSE = 42$$

$$SSTO := \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \left[\left(Y_{blocks_{i,j}}\right)_k - \mu_{dotdot}\right]^2 \qquad SSTO = 766$$

Vector form:

$b \cdot n \cdot \left(\alpha^T \cdot \alpha\right) = (156.5)$

$\beta := \text{stack}\left(\beta_{BinA}^{\langle 1 \rangle}, \beta_{BinA}^{\langle 2 \rangle}\right)$

$$\beta = \begin{pmatrix} 7.25 \\ -5.75 \\ 7.5 \\ -7.25 \\ 5.75 \\ -7.5 \end{pmatrix}$$

β in vector form ^

$n \cdot \beta^T \cdot \beta = (567.5)$

$e^T \cdot e = (42)$

$\left(Y - \mu_{dotdot}\right)^T \cdot \left(Y - \mu_{dotdot}\right) = (766)$

ANOVA Table for Nested design:

Sum of Squares: Degrees of Freedom: Mean Squares:

$SSA = 156.5$	$df_A := a - 1$	$df_A = 2$	$MSA := \frac{SSA}{df_A}$	$MSA = 78.25$
$SSB_{inA} = 567.5$	$df_{BinA} := a \cdot (b - 1)$	$df_{BinA} = 3$	$MSB_{inA} := \frac{SSB_{inA}}{df_{BinA}}$	$MSB_{inA} = 189.1667$
$SSE = 42$	$df_E := a \cdot b \cdot (n - 1)$	$df_E = 6$	$MSE := \frac{SSE}{df_E}$	$MSE = 7$
$SSTO = 766$	$df_T := N - 1$	$df_T = 11$		

^ values verified KNNL p. 1097

F-Test for Independent Factor Effects:

Hypotheses:

H_0: $\alpha_i = 0$ ***for all*** **i**

H_1: ***At least one*** α_i **<> 0**

Test Statistic:

$F := \frac{MSA}{MSE}$ $F = 11.1786$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$CV := qF\left[1 - \alpha, (a - 1), (n - 1) \cdot a \cdot b\right]$ $CV = 5.1433$

$a - 1 = 2$

$(n - 1) \cdot a \cdot b = 6$

Decision Rule:

IF F > C, THEN REJECT H_0 OTHERWISE ACCEPT H_0

Probability Value:

< test verified KNNL p. 1098

$P := 1 - pF[F, a - 1, (n - 1) \cdot a \cdot b]$ $P = 0.0095$

F-Test for Nested Factor Specific Effects:

Hypotheses:

H_0: $\beta_{j\,in\,i}$ **= 0** ***for all values of j within a specific level i***

H_1: ***At least one*** $\beta_{j\,in\,i}$ **<> 0**

Test Statistic:

$F := \frac{MSB_{inA}}{MSE}$ $F = 27.0238$

Critical Value of the Test:

$\alpha := 0.05$ **< Probability of Type I error must be explicitly set**

$CV := qF\left[1 - \alpha, a \cdot (b - 1), (n - 1) \cdot a \cdot b\right]$ $CV = 4.7571$

Decision Rule:

IF F > C, THEN REJECT H_0 OTHERWISE ACCEPT H_0

Probability Value:

< test verified KNNL p. 1098

$P := 1 - pF[F, a \cdot (b - 1), (n - 1) \cdot a \cdot b]$ $P = 0.0007$

Prototype in R:

```
#NESTED ANOVA
require(car)

#READ STRUCTURED DATA TABLE WITH NUMERIC CODED FACTOR
K=read.table("c:/2008LinearModelsData/TrainingSchoolR.txt")
K
attach(K)
Y=Score
A=factor(School)
B=factor(Instructor)

contrasts(A)=contr.sum
contrasts(B)=contr.sum

FM=lm(Y~A+B%in%A)
anova(FM)
Anova(FM, type="3")
detach(K)
```

```
> K
   Score School Instructor
1     25      1          1
2     29      1          1
3     14      1          2
4     11      1          2
5     11      2          1
6      6      2          1
7     22      2          2
8     18      2          2
9     17      3          1
10    20      3          1
11     5      3          2
```

```
> anova(FM)
Analysis of Variance Table

Response: Y
          Df Sum Sq Mean Sq F value   Pr(>F)
A          2 156.50   78.25  11.179 0.009473 **
A:B        3 567.50  189.17  27.024 0.000697 ***
Residuals  6  42.00    7.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA tables are identical >

```
> Anova(FM, type="3")
Anova Table (Type III tests)

Response: Y
            Sum Sq Df F value    Pr(>F)
(Intercept) 2700.0  1 385.714 1.130e-06 ***
A            156.5  2  11.179  0.009473 **
A:B          567.5  3  27.024  0.000697 ***
Residuals     42.0  6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

factor A results confirmed >

factor B in A results confirmed >

KNNL p. 1098