2-Factor Randomized Complete Block ANOVA

Engineering Statistics II
Section 11.1

Josh Engwer

TTU

2018

PART I:

Nuisance Factors

Randomized Blocking

Nuisance Factors (Definition & Types)

Alas, undesirable factors may affect an experiment:

Definition

(Nuisance Factor)

Given a designed experiment.
An uninteresting factor that may affect the response is a **nuisance factor**.

The three types of nuisance factors are dealt with via different techniques‡:

NUISANCE FACTOR TYPE:	MITIGATION:	COVERED IN THIS COURSE?
Unknown & Uncontrollable	Randomization & Double-Blinding	Randomization: Yes Double-Blinding: No
Known & Uncontrollable	Analysis of Covariance (ANCOVA)	No
Known & Controllable	**Randomized Blocking**	**Yes (Next slides)**

"Block what you can, randomize what you cannot." – G.E.P. Box, 1978

‡D.C. Montgomery, *Design & Analysis of Experiments*, 7^{th} Ed, Wiley, 2009. (§4.1)

Nuisance Factors (Examples)

Alas, undesirable factors may affect an experiment‡:

NUISANCE FACTOR TYPE:	EXAMPLES:
Unknown & Uncontrollable	Bias of Designer(s) of Exp. Bias of Administrator(s) of Exp. Bias of Human Subject(s) in Exp.
Known & Uncontrollable	Outside Weather (temp, humidity, wind, ...) Ambient Temp. in Large Warehouse Life Experience of Human Subjects
Known & Controllable	**Origin/Purity of Raw Material Batches** **Ambient Temp. in Small Room** **Accuracy/Precision of Workers** **Accuracy/Precision of Machines** **Time of Day when Exp. is Conducted** **Manufacturer of Comparable Tools** **Age Group of Human Subjects**

"Block what you can, randomize what you cannot." – G.E.P. Box, 1978

‡D.C. Montgomery, *Design & Analysis of Experiments*, 7^{th} Ed, Wiley, 2009. (§4.1)

Nuisance Factors (Biases)

Bias in an experiment is minimized via Randomization & Double-Blinding:

NUISANCE FACTOR TYPE:	**EXAMPLES:**
Unknown & Uncontrollable	Bias of Designer(s) of Exp. Bias of Administrator(s) of Exp. Bias of Human Subject(s) in Exp.

Since randomization of experiments encountered in this course is always assumed, we will not delve into the details of bias.

However, interested readers may consult the following papers:

H. Malone, H. Nicholl, C. Tracey, "Awareness and Minimisation of Systematic Bias in Research", *British Journal of Nursing*, **23** (2014), 279-282.

D. Sackett, "Bias in Analytic Research", *J. Chronic Diseases*, **32** (1979), 51-63.

A.-M. Šimundić, "Bias in Research", *Biochemia Medica*, **23** (2013), 12-15.

Randomized Blocking (Motivation)

Randomized blocking[♠][♣] controls for the presence of a nuisance factor.

1-Factor Complete Randomization:

- Collect experimental units (EU's).
- Randomly assign the EU's to the levels of Factor A.
- Measure each EU appropriately.

2-Factor Complete Randomization:

- Collect experimental units (EU's).
- Randomly assign the EU's to the level combinations of Factors A & B.
- Measure each EU appropriately.

2-Factor Randomized Blocking:

- Collect experimental units (EU's).
- Identify relevant nuisance factor & call it **Block Factor B**.
- Determine the nuisance factor level of each EU.
- Assign all the EU's with equal nuisance factor level to its own **block**.
- <u>Within each block</u>, randomly assign EU's to levels of Factor A.
- Measure each EU appropriately.

[♠] R.A. Fisher, "The Arrangement of Field Experiments", *J. Ministry Agr.*, **33** (1926), 503-513.

[♣] R.A. Fisher, *The Design of Experiments*, Oliver & Boyd, 1935. (Ch4)

PART II:
2-FACTOR RANDOMIZED COMPLETE BLOCK ANOVA
(2F rcbANOVA)

2-Factor Randomized Complete Block Design

An example **randomized complete block design** entails the following:

- Collect 6 relevant EU's: $\text{EU}_1, \text{EU}_2, \text{EU}_3, \text{EU}_4, \text{EU}_5, \text{EU}_6$
- Determine EUs' nuisance levels (which is in parentheses):
  $\text{EU}_{1(3)}, \text{EU}_{2(1)}, \text{EU}_{3(1)}, \text{EU}_{4(2)}, \text{EU}_{5(3)}, \text{EU}_{6(2)}$
- Produce a random shuffle sequence for each nuisance level:
  Lvl 1: $(3;\ 2)$, Lvl 2: $(4;\ 6)$, Lvl 3: $(5;\ 1)$
- Use random shuffle sequence to assign the EU's into the 6 groups:

BLOCK B: $\rightarrow$ / **FACTOR A:** $\downarrow$	Level 1	Level 2	Level 3
Level 1	$\text{EU}_{3(1)}$	$\text{EU}_{4(2)}$	$\text{EU}_{5(3)}$
Level 2	$\text{EU}_{2(1)}$	$\text{EU}_{6(2)}$	$\text{EU}_{1(3)}$

- Measure each EU appropriately (note the change in notation):

BLOCK B: $\rightarrow$ / **FACTOR A:** $\downarrow$	Level 1 $(x_{\bullet 1})$	Level 2 $(x_{\bullet 2})$	Level 3 $(x_{\bullet 3})$
Level 1 $(x_{1\bullet})$	x_{11}	x_{12}	x_{13}
Level 2 $(x_{2\bullet})$	x_{21}	x_{22}	x_{23}

2F rcbANOVA Fixed Effects Model Assumptions

Proposition

- *(**1 Desired Factor**) The sole factor of interest has I levels.*
- *(**1 Nuisance Factor**) The sole nuisance factor has J levels.*
- *(**All Factor Levels are Considered**) AKA Fixed Effects.*
- *(**1 Measurement per Group**) Each of the IJ groups has one exp unit.*

- *(**Random Assignment within Blocks**)*

such that (s.t.)

- *(**Nuisance Same in Block**) Within block, nearly same nuisance values.*
- *(**Nuisance Differs across Blocks**) Blocks differ by nuisance value.*

- *(**Independence**) All measurements on units are independent.*
- *(**Normality**) All IJ groups are approximately normally distributed.*
- *(**Equal Variances**) All IJ groups have approximately same variance.*
- *(**Factor and Block are not Interactive**)*

1DF 1NF AFLaC 1MpG | RAwB s.t. NSiB NDaB | I.N.EV FaBanI

2-Factor rcbANOVA Linear Model (Fixed Effects)

2F rcbANOVA Fixed Effects Linear Model		
(I,J)	$\equiv$	(# levels of factor A, # levels of blocked nuisance factor B)
X_{ij}	$\equiv$	rv for observation at (i,j)-level of (factor A, block B)
μ	$\equiv$	Mean avg response over all levels of (factor A, block B)
$(\alpha_i^A, \alpha_j^{[B]})$	$\equiv$	(Effect of i^{th}-level factor A, Effect of j^{th}-level block B)
E_{ij}	$\equiv$	Deviation from μ due to random error

ASSUMPTIONS: $E_{ij} \overset{iid}{\sim}$ Normal $\left(0, \sigma^2\right)$

$$X_{ij} = \mu + \alpha_i^A + \alpha_j^{[B]} + E_{ij} \quad \text{where} \quad \sum_i \alpha_i^A = \sum_j \alpha_j^{[B]} = 0$$

$$H_0^A : \quad \text{All} \quad \alpha_i^A = 0$$
$$H_A^A : \quad \text{Some} \quad \alpha_i^A \neq 0$$

$X_{ij} \overset{IND}{\sim} \ldots \equiv$ rv's X_{ij} are independently distributed as ...

$E_{ij} \overset{iid}{\sim} \ldots \equiv$ rv's E_{ij} are independently and identically distributed as ...

Sums of Squares as a “Partitioning” of Variation Explanation for 2F rcbANOVA

$$\underbrace{\text{SS}_{total}}_{\substack{Total\ Variation\\ in\ Experiment}} = \underbrace{\text{SS}_{A}}_{\substack{Variation\ due\\ to\ Factor\ A}} + \underbrace{\text{SS}_{[B]}}_{\substack{Variation\ due\\ to\ Blocks\ B}} + \underbrace{\text{SS}_{res}}_{\substack{Unexplained\\ Variation}}$$

$$\sum_{ij}(x_{ij} - \hat{\mu})^2 = \sum_{ij}(\hat{\alpha}_i^A)^2 + \sum_{ij}\left(\hat{\alpha}_j^{[B]}\right)^2 + \sum_{ij}(x_{ij}^{res})^2$$

$$\underbrace{\nu}_{\substack{Total\ dof's\ in\\ Experiment}} = \underbrace{\nu_A}_{\substack{Factor\ A\\ dof's}} + \underbrace{\nu_{[B]}}_{\substack{Blocks\ B\\ dof's}} + \underbrace{\nu_{res}}_{\substack{'Within\ Blocks'\\ dof's}}$$

$$\nu = IJ - 1, \quad \nu_A = I - 1, \quad \nu_{[B]} = J - 1, \quad \nu_{res} = (I-1)(J-1)$$

Remember, randomization occurs only <u>within blocks</u>‡.

‡D.C. Montgomery, *Design & Analysis of Experiments*, 7^{th} Ed, Wiley, 2009. (§4.1)

2F rcbANOVA F-Test

1. $\nu_A = I - 1,\ \nu_{[B]} = J - 1,\ \nu_{res} = (I-1)(J-1)$

2. $\bar{x}_{i\bullet} := \frac{1}{J}\sum_j x_{ij}, \quad \bar{x}_{\bullet j} := \frac{1}{I}\sum_i x_{ij}$
3. $\bar{x}_{\bullet\bullet} := \frac{1}{IJ}\sum_i \sum_j x_{ij}$

4. $$\begin{cases} \text{SS}_{res} := \sum_{ij}(x^{res}_{ijk})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_{i\bullet} - \bar{x}_{\bullet j} + \bar{x}_{\bullet\bullet})^2 \\ \text{SS}_A := \sum_{ij}(\hat{\alpha}^A_i)^2 = \sum_i \sum_j (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2 \\ \text{SS}_{[B]} := \sum_{ij}(\hat{\alpha}^{[B]}_j)^2 = \sum_i \sum_j (\bar{x}_{\bullet j} - \bar{x}_{\bullet\bullet})^2 \end{cases}$$

(Optional) $\text{SS}_{total} := \sum_{ij}(x_{ij} - \hat{\mu})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_{\bullet\bullet})^2$

2F rcbANOVA F-Test

5. $$\text{MS}_A = \frac{\text{SS}_A}{\nu_A}, \quad \text{MS}_{[B]} = \frac{\text{SS}_{[B]}}{\nu_{[B]}}, \quad \text{MS}_{res} = \frac{\text{SS}_{res}}{\nu_{res}}$$

6. $$f_A = \frac{\text{MS}_A}{\text{MS}_{res}}, \quad f_{[B]} = \frac{\text{MS}_{[B]}}{\text{MS}_{res}}$$

7. $$\begin{cases} p_A & := & \mathbb{P}(F > f_A) & \approx & 1 - \Phi_F(f_A; \nu_A, \nu_{res}) \\ p_{[B]} & := & \mathbb{P}(F > f_{[B]}) & \approx & 1 - \Phi_F(f_{[B]}; \nu_{[B]}, \nu_{res}) \end{cases}$$

8. $$\begin{cases} \text{If} \quad p_A \leq \alpha & \text{then reject } H_0^A \text{, else accept } H_0^A. \\ \text{If} \quad p_{[B]} \leq \alpha & \text{then the blocking reduced } MS_{res} \text{ vs. 1F ANOVA.} \end{cases}$$

2F rcbANOVA Table

2-Factor rcbANOVA Table (Significance Level α)						
Variation Source	**df**	**Sum of Squares**	**Mean Square**	**F Stat Value**	**P-value**	**Decision**
Factor A	ν_A	SS_A	MS_A	f_A	p_A	Acc/Rej H_0^A
Blocks B	$\nu_{[B]}$	$\text{SS}_{[B]}$	$\text{MS}_{[B]}$	$f_{[B]}$	$p_{[B]}$	$*$
Error	ν_{res}	SS_{res}	MS_{res}			
Total	ν	SS_{total}				

*Computing $\text{SS}_{[B]}, \text{MS}_{[B]}, f_{[B]}, p_{[B]}$ is optional but recommended as $p_{[B]} \leq \alpha$ implies that the blocking choice results in a significantly smaller MS_{res} than using 1F bcrANOVA, thus the blocked nuisance factor has a significant effect.

On the other hand, if $p_{[B]} > \alpha$, then the particular blocking is not beneficial. The remedy is to block on a (hopefully) more relevant nuisance factor.

2F rcbANOVA (Expected Mean Squares)

Proposition

Given 2-factor experiment satisfying the 2F rcbANOVA assumptions. Then:

$$\begin{aligned}
\mathbb{E}[MS_{res}] &= \sigma^2 \\
\mathbb{E}[MS_A] &= \sigma^2 + \frac{J}{I-1}\sum_i (\alpha_i^A)^2 \\
\mathbb{E}[MS_{[B]}] &= \sigma^2 + \frac{I}{J-1}\sum_j \left(\alpha_j^{[B]}\right)^2
\end{aligned}$$

PROOF: Omitted as it's similar (and simpler) to 2F bcrANOVA.

Mean Squares as Point Estimators of σ^2

Proposition

(Point Estimation of Mean Squares)
Given a 2F balanced exp. satisfying the 2F rcbANOVA assumptions. Then:

(i) *Regardless of the truthness of* $H_0^A, H_0^{[B]}$ $\implies$ $\mathbb{E}[MS_{res}] = \sigma^2$

(ii) H_0^A *is true* $\implies$ $\mathbb{E}[MS_A] = \sigma^2$, H_0^A *is false* $\implies$ $\mathbb{E}[MS_A] > \sigma^2$

PROOF: Omitted as it's similar (and simpler) to 2F bcrANOVA.

YEAR	NAME	MEASURE
1925[†]	Fisher (Eta-Squared)	$\hat{\eta}^2_A := \frac{SS_A}{SS_{total}} = \frac{\nu_A f_A}{\nu_A f_A + \nu_{[B]} f_{[B]} + \nu_{res}}$ $\hat{\eta}^2_{[B]} := \frac{SS_{[B]}}{SS_{total}} = \frac{\nu_{[B]} f_{[B]}}{\nu_A f_A + \nu_{[B]} f_{[B]} + \nu_{res}}$ $\hat{\eta}^2_{res} := \frac{SS_{res}}{SS_{total}}$
1965[‡]	Cohen[GW],[LH] (Partial η^2)	$\hat{\eta}^2_{(A)} := \frac{SS_A}{SS_A + SS_{res}} = \frac{\nu_A f_A}{\nu_A f_A + \nu_{res}}$ $\hat{\eta}^2_{([B])} := \frac{SS_{[B]}}{SS_{[B]} + SS_{res}} = \frac{\nu_{[B]} f_{[B]}}{\nu_{[B]} f_{[B]} + \nu_{res}}$

$\hat{\eta}^2_A + \hat{\eta}^2_{[B]} + \hat{\eta}^2_{res} = 1$ but $\hat{\eta}^2_{(A)} + \hat{\eta}^2_{([B])} > 1$

[†] R.A. Fisher, *Statistical Methods for Reasearch Workers*, 1925.

[‡] B.B. Wolman (Ed.), *Handbook of Clinical Psychology*, 1965. (§5 by J. Cohen)

EFFECT SIZE VALUE:	INTERPRETATION:
$\hat{\eta}^2_A := \frac{SS_A}{SS_A + SS_{[B]} + SS_{res}} = 0.38$	38% of the variation in the reponse is due to Factor A
$\hat{\eta}^2_{[B]} := \frac{SS_{[B]}}{SS_A + SS_{[B]} + SS_{res}} = 0.02$	2% of the variation in the reponse is due to Block B
$\hat{\eta}^2_{res} := \frac{SS_{res}}{SS_A + SS_{[B]} + SS_{res}} = 0.60$	60% of the variation in the reponse is unexplained with experiment
$\hat{\eta}^2_{(A)} := \frac{SS_A}{SS_A + SS_{res}} = 0.43$	43% of the variation possibly due to Factor A is actually due to Factor A
$\hat{\eta}^2_{([B])} := \frac{SS_{[B]}}{SS_{[B]} + SS_{res}} = 0.72$	72% of the variation possibly due to Block B is actually due to Block B

YEAR	NAME	MEASURE
1963[†]	Hays (Omega-Squared)	$\hat{\omega}^2_A := \frac{SS_A - \nu_A MS_{res}}{SS_{total} + MS_{res}} = \frac{\nu_A f_A - \nu_A}{\nu_A f_A + \nu_{[B]} f_{[B]} + n}$ $\hat{\omega}^2_{[B]} := \frac{SS_{[B]} - \nu_{[B]} MS_{res}}{SS_{total} + MS_{res}} = \frac{\nu_{[B]} f_{[B]} - \nu_{[B]}}{\nu_A f_A + \nu_{[B]} f_{[B]} + n}$
1979[‡]	Keren-Lewis (Partial ω^2)	$\hat{\omega}^2_{(A)} := \frac{SS_A - \nu_A MS_{res}}{SS_A + (n - \nu_A) MS_{res}} = \frac{\nu_A (f_A - 1)}{\nu_A (f_A - 1) + n}$ $\hat{\omega}^2_{([B])} := \frac{SS_{[B]} - \nu_{[B]} MS_{res}}{SS_{[B]} + (n - \nu_{[B]}) MS_{res}} = \frac{\nu_{[B]} (f_{[B]} - 1)}{\nu_{[B]} (f_{[B]} - 1) + n}$

$$n := IJ = (1 + \nu_A)(1 + \nu_{[B]})$$

† W.L. Hays, *Statistics for Psychologists*, 1963.

‡ G. Keren, C. Lewis, “Partial Omega Squared for ANOVA Designs”, *Educational & Psychological Measurement*, **39** (1979), 119-128.

2F rcbANOVA Tukey Post-Hoc Comparisons

Suppose a 2-Factor rcbANOVA results in the rejection of H_0^A.
Then, at least two of the pop. means significantly differ, but which ones?

Proposition

Given a 2-factor experiment with I levels of factor A and J levels of blocked nuisance factor B where 2F rcbANOVA rejects H_0^A at significance level α. Then, to find which levels of factor A significantly differ:

1. *Compute the factor A significant difference width:* $[\nu_{res} := (I-1)(J-1)]$

$$w_A = q^*_{I,\nu_{res};\alpha} \cdot \sqrt{MS_{res}/J}$$

2. *Sort the I factor A level means in ascending order:*

$$\bar{x}_{(1)\bullet} \leq \bar{x}_{(2)\bullet} \leq \cdots \leq \bar{x}_{(I)\bullet}$$

3. *For each sorted group mean $\bar{x}_{(i)\bullet}$:*
   - *If $\bar{x}_{(i+1)\bullet} \notin \left[\bar{x}_{(i)\bullet},\ \bar{x}_{(i)\bullet} + w_A\right]$, repeat STEP 3 with next sorted mean.*
   - *Else, underline $\bar{x}_{(i)\bullet}$ and all larger means within a distance of w_A w/ new line.*

NOTE: Tukey Post-Hoc Comp's are used only for factor A, not for block B.

Textbook Logistics for Section 11.1

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Sum of Squares of Factor A	SSTr	SS_A
Mean Square of Factor A	MSTr	MS_A
Sum of Squares of Error	SSE	SS_{res}
Mean Square of Error	MSE	MS_{res}
Effect of i^{th} Factor A	α_i	α_i^A
Null Hypothesis for Factor A	H_{0A}	H_0^A
Alt. Hypothesis for Factor A	H_{aA}	H_A^A
Expected Value	$E(X)$	$\mathbb{E}[X]$
Variance	$V(X)$	$\mathbb{V}[X]$

- Ignore "Models for Random Effects" section.
  - The ANOVA procedure is identical as for fixed effects linear models.
  - However, model assumption checking is subtler and trickier.

Fin.