NUISANCE FACTORS [DEVORE 11.1]

NUISANCE FACTORS (DEF'N): An uninteresting factor that may affect the response is a **nuisance factor**.

NUISANCE FACTORS (TYPES): 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**

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)

2-FACTOR RANDOMIZED COMPLETE BLOCK ANOVA (2F rcbANOVA) [DEVORE 11.1]

- **2F rcbANOVA (RANDOMIZED COMPLETE BLOCK DESIGN):** As an example:
  - Collect 6 relevant experimental units (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.
  - Measure each EU appropriately. (note the change in notation)

BLOCK B: → / FACTOR A: ↓	Lvl 1	Lvl 2	Lvl 3
Level 1	$\text{EU}_{3(1)}$	$\text{EU}_{4(2)}$	$\text{EU}_{5(3)}$
Level 2	$\text{EU}_{2(1)}$	$\text{EU}_{6(2)}$	$\text{EU}_{1(3)}$

$\overset{MEASURE}{\Longrightarrow}$

B: → / A: ↓	Lvl 1 ($x_{\bullet 1}$)	Lvl 2 ($x_{\bullet 2}$)	Lvl 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):**
  - ⋆ (**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**)

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

- **2F rcbANOVA (SUMS OF SQUARES "PARTITION" VARIATION):**

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

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

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

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

- **2F rcbANOVA (EXPECTED MEAN SQUARES):**

$$(i)\ \mathbb{E}[\text{MS}_{res}] = \sigma^2 \qquad (ii)\ \mathbb{E}[\text{MS}_A] = \sigma^2 + \frac{J}{I-1}\sum_i(\alpha_i^A)^2 \qquad (iii)\ \mathbb{E}[\text{MS}_{[B]}] = \sigma^2 + \frac{I}{J-1}\sum_j\left(\alpha_j^{[B]}\right)^2$$

- **2F rcbANOVA (POINT ESTIMATORS OF σ^2):**

$$(i)\quad \text{Regardless of the truthfulness of } H_0^A, H_0^{[B]} \implies \mathbb{E}[\text{MS}_{res}] = \sigma^2$$

$$(ii)\quad H_0^A \text{ is true} \implies \mathbb{E}[\text{MS}_A] = \sigma^2, \quad H_0^A \text{ is false} \implies \mathbb{E}[\text{MS}_A] > \sigma^2$$

2-FACTOR RANDOMIZED COMPLETE BLOCK ANOVA (2F rcbANOVA) [DEVORE 11.1]

- **2F rcbANOVA (FIXED EFFECTS LINEAR MODEL):**

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: $E_{ij} \overset{iid}{\sim} \text{Normal}\left(0, \sigma^2\right)$
		$X_{ij} = \mu + \alpha_i^A + \alpha_j^{[B]} + E_{ij}$ where $\sum_i \alpha_i^A = \sum_j \alpha_j^{[B]} = 0$
		H_0^A : All $\alpha_i^A = 0$
		H_A^A : Some $\alpha_i^A \neq 0$

- **2F rcbANOVA (F-TEST PROCEDURE):**

1. Determine df's: $\nu_A = I - 1,\ \nu_{[B]} = J - 1,\ \nu_{res} = (I-1)(J-1)$
2. Compute Group Means (if not provided): $\overline{x}_{i\bullet} := \frac{1}{J}\sum_j x_{ij}, \quad \overline{x}_{\bullet j} := \frac{1}{I}\sum_i x_{ij}$
3. Compute Grand Mean: $\overline{x}_{\bullet\bullet} := \frac{1}{IJ}\sum_i \sum_j x_{ij}$
4. Compute $\text{SS}_{res} := \sum_{ij}(x_{ij}^{res})^2 = \sum_i \sum_j (x_{ij} - \overline{x}_{i\bullet} - \overline{x}_{\bullet j} + \overline{x}_{\bullet\bullet})^2$
5. Compute $\text{SS}_A \ := \sum_{ij}(\hat{\alpha}_i^A)^2 = \sum_i \sum_j (\overline{x}_{i\bullet} - \overline{x}_{\bullet\bullet})^2$
6. Compute $\text{SS}_{[B]} := \sum_{ij}(\hat{\alpha}_j^{[B]})^2 = \sum_i \sum_j (\overline{x}_{\bullet j} - \overline{x}_{\bullet\bullet})^2$

   (Optional) $\text{SS}_{total} := \sum_{ij}(x_{ij} - \hat{\mu})^2 = \sum_i \sum_j (x_{ij} - \overline{x}_{\bullet\bullet})^2$
7. Compute Mean Squares: $\text{MS}_{res} := \dfrac{\text{SS}_{res}}{\nu_{res}}, \qquad \text{MS}_A := \dfrac{\text{SS}_A}{\nu_A}, \qquad \text{MS}_{[B]} = \dfrac{\text{SS}_{[B]}}{\nu_{[B]}}$
8. Compute Test Statistic Value(s): $f_A = \dfrac{\text{MS}_A}{\text{MS}_{res}}, \quad f_{[B]} = \dfrac{\text{MS}_{[B]}}{\text{MS}_{res}}$
9. If using software, compute P-value(s): $\begin{cases} p_A & := & \mathbb{P}(F > f_A) & \approx & 1 - \Phi_F(f_A; \nu_A, \nu_{res}) \\ p_{[B]} & := & \mathbb{P}(F > f_{[B]}) & \approx & 1 - \Phi_F(f_{[B]}; \nu_{[B]}, \nu_{res}) \end{cases}$
10. Render Decision(s): $\begin{cases} \text{If} \quad p_A \leq \alpha \quad \text{or} \quad f_A > f^*_{\nu_A, \nu_{res}; \alpha} & \text{then reject } H_0^A\text{, else accept } H_0^A. \\ \text{If} \quad p_{[B]} \leq \alpha \quad \text{or} \quad f_{[B]} > f^*_{\nu_{[B]}, \nu_{res}; \alpha} & \text{then the blocking reduced MS}_{res} \text{ vs. 1F ANOVA.} \end{cases}$

- **2F rcbANOVA (SUMMARY 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 $p_{[B]} \leq \alpha$ or $f_{[B]} > f^*_{\nu_{[B]}, \nu_{res}; \alpha}$ 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 $p_{[B]} > \alpha$ or $f_{[B]} < f^*_{\nu_{[B]}, \nu_{res}; \alpha}$, then the particular blocking is not beneficial.
The remedy is to block on a (hopefully) more relevant nuisance factor.

2F rcbANOVA (EFFECT SIZES & POST-HOC TESTS) [DEVORE 11.1]

- **2F rcbANOVA (EFFECT SIZE MEASURES & THEIR INTERPRETATIONS):**

YEAR	NAME	EFFECT SIZE VALUE:	INTERPRETATION:
1925†	Fisher (Eta-Squared)	$\hat{\eta}^2_A := \frac{SS_A}{SS_A+SS_{[B]}+SS_{res}} = 0.38$	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
1965‡	Cohen (Partial η^2)	$\hat{\eta}^2_{(A)} := \frac{SS_A}{SS_A+SS_{res}} = 0.43$	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

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

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

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

- **2F rcbANOVA (MORE EFFECT SIZE MEASURES):**

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", *Edu. & Psych. 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?
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:

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

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

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

EX 11.1.1: The lifetimes of three light bulb brands were blocked by raw material batch and then measured:

BULB LIFETIME (in years)						
BLOCK B: $\rightarrow$ **FACTOR A:** $\downarrow$	Batch 1 ($x_{\bullet 1}$)	Batch 2 ($x_{\bullet 2}$)	Batch 3 ($x_{\bullet 3}$)	Batch 4 ($x_{\bullet 4}$)	Batch 5 ($x_{\bullet 5}$)	**TOTAL** ($\sum_j x_{ij}$)
Brand 1 ($x_{1\bullet}$)	9.22	9.07	8.95	8.98	9.54	**45.76**
Brand 2 ($x_{2\bullet}$)	8.92	8.88	9.10	8.71	8.85	**44.46**
Brand 3 ($x_{3\bullet}$)	9.08	8.99	9.06	8.93	9.02	**45.08**
TOTAL ($\sum_i x_{ij}$)	**27.22**	**26.94**	**27.11**	**26.62**	**27.41**	$\sum_i \sum_j x_{ij} =$ **135.30**

(a) Formulate this experiment as a 2-Factor fixed effects linear model. In this context, what does "fixed effects" assume?

(b) State the appropriate null hypothesis H_0^A and alternative hypothesis H_A^A.

(c) Perform a 2-Factor Randomized Complete Block ANOVA (2F rcbANOVA) with ($\alpha = 0.01$) significance level.

Was the chosen blocking effective? To save time and tedium: $\text{SS}_{total} = 0.4946$, $\text{SS}_{res} \approx 0.20595$

(d) Compute & interpret the eta-squared and partial eta-squared effect size measures: $\hat{\eta}^2_A, \hat{\eta}^2_{[B]}$; $\hat{\eta}^2_{(A)}, \hat{\eta}^2_{([B])}$

EX 11.1.2: The counts of M&M's® peanut chocolate candies in seven equal-size bags are provided in this table†:

NUMBER OF CHOCOLATE CANDIES OF A GIVEN COLOR IN A GIVEN BAG								
BLOCK B: → **FACTOR A:** ↓	Bag 1 ($x_{\bullet 1}$)	Bag 2 ($x_{\bullet 2}$)	Bag 3 ($x_{\bullet 3}$)	Bag 4 ($x_{\bullet 4}$)	Bag 5 ($x_{\bullet 5}$)	Bag 6 ($x_{\bullet 6}$)	Bag 7 ($x_{\bullet 7}$)	**TOTAL** ($\sum_j x_{ij}$)
Blue ($x_{1\bullet}$)	8	7	5	7	6	8	6	**47**
Red ($x_{2\bullet}$)	2	2	5	3	5	4	5	**26**
Orange ($x_{3\bullet}$)	1	0	0	1	1	2	1	**6**
Green ($x_{4\bullet}$)	0	1	0	2	0	3	2	**8**
Brown ($x_{5\bullet}$)	5	6	6	7	5	7	5	**41**
Yellow ($x_{6\bullet}$)	2	1	3	1	2	3	1	**13**
TOTAL ($\sum_i x_{ij}$)	**18**	**17**	**19**	**21**	**19**	**27**	**20**	$\sum_i \sum_j x_{ij} =$ **141**

†This table is a simplified and modified version of the table (and experiment) found in:

T. Lin, M.S. Sanders, "A Sweet Way to Learn DoE", *Quality Progress*, **39** (2006), 88.

(a) Formulate this experiment as a 2-Factor fixed effects linear model.

(b) State the appropriate null hypothesis H_0^A and alternative hypothesis H_A^A.

(c) Perform a 2-Factor Randomized Complete Block ANOVA (2F rcbANOVA) with ($\alpha = 0.05$) significance level.

Was the blocking effective? To save time: $\text{SS}_{total} \approx 257.643$, $\text{SS}_A \approx 217.357$, $\text{SS}_{[B]} \approx 10.810$, $\text{SS}_{res} \approx 29.476$

(d) Compute & interpret the eta-squared and partial eta-squared effect size measures: $\hat{\eta}_A^2$; $\hat{\eta}_{(A)}^2$

(e) Perform the appropriate Tukey Complete Pairwise Post-Hoc Comparison.