UPM
UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI

FAKULTI PERTANIAN
FACULTY OF AGRICULTURE
فاكولتي ڤرتانين

AGR5201
Advanced Statistical Methods

Nested design
(One and two factors)

Topic outline

1.0 Introduction

1.1 Factors and effects

1.2 ANOVA of nested design

2.0 Nested design with fixed and random effects

2.1 One factor nested design – fixed model

2.2 Two factor nested design – fixed model

2.3 Two factor nested design – mixed model

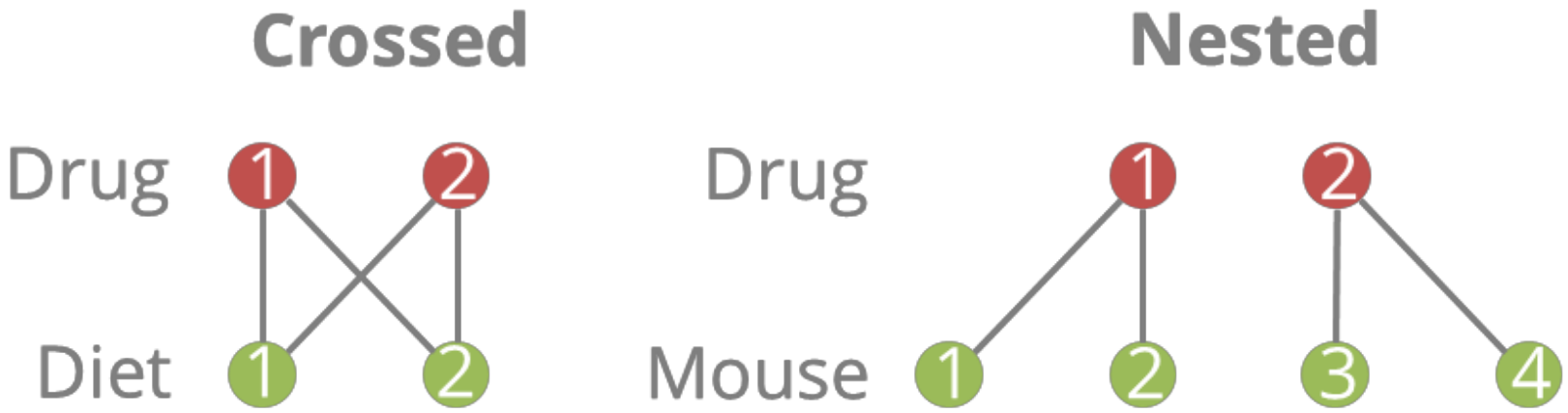
3.0 Difference between Nested and Split plot

Reference website:

https://online.stat.psu.edu/stat503/lesson/14/14.1

1.0 Introduction

Experiment type: Crossed vs. Nested



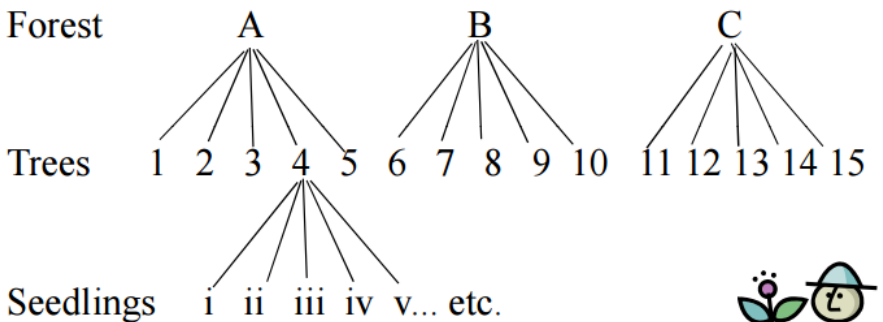
All possible combinations of drug and diet

Each unit of mouse is unique to each unit of drug type

1.0 Introduction

Nested design

- A nested design (hierarchical design) is used for experiments when the experimental units are sub-sampled
- Example: a forest geneticist collects 5 seeds from 5 trees in each of 3 forests. The design looks like this:



Total of 75 seedlings.

1.0 Introduction

When to use nested design?

- In general, nested design is used when the treatment (one factor) or one of factors (factorial) is not possible to be randomized or sometimes not replicated (depends on nature of factor).

nested desin is used when one factor or one of factors(factorial) is not possible to be randomized or sometimes not replicated

- *One-factor experiment*
  - The assignment of treatment are not randomized due to constraints in experimental setup.

- *Factorial Experiments (Factor A and B)*
  - With "Fixed Effects" and "Random Effects"
  - The levels of factor (B) are not identical to each other at different levels of factor (A), although they might have the same labels.
  - For example, if A is school and B is teacher, teacher 1 will differ between the schools.

1.1 Factors and effects

Number of factors

One factor (one-stage)

Due to some constrains in the experimental procedure, the treatment is not replicated, but the experimental units or replications are randomly arranged within each treatment

Two factors - Factor A and B (two-stage)

Factor A (big factor) is not replicated/randomized
Only factor B (nested factor) is replicated and randomly arranged within Factor A (big factor)

1.1 Factors and effects

Fixed vs. Random effects

Fixed effects	Random effects

✓ The effect of factor are **FIXED** if you choose the specific level and intrinsic value. So:
• The inference will be limited exclusively to the specific levels that appear in the experiment
• The levels may be repeated in the future
  • specific recommendation for agricultural practices or future experiments
• The statistics used to summarize the results:
  • means, least square means, treatment effects
 	
if FIXED, interested in means
	✓ The effects of a factor are assumed to be **RANDOM** if you chose them to represent a larger reference population of levels.
	✓ Obtained through sampling → random sample
	• The inferences from this experiment applied to the reference population from which they were drawn.
	• The specific levels may not be repeated in the future nor the specific recommendations be made regarding specific levels
	• The statistics: variances, correlations and/or regression statistics
	
	if RANDOM, interested in variances

1.1 Factors and effects

Two-stage nested design

- ***Fixed model***
  - Both big and small factors are treated as fixed effects
  - Interested in the mean for each factor level
- ***Mixed model***
  - Big factor is treated as fixed effect → interested in mean
  - Small factor (nested factor) is treated as random → not interested in the mean, but the variance.

1.2 ANOVA of nested design

Types of nested design

- The nested factors can be treated as fixed or random
- This types of nested factors will determine the error term in the F test for the big plot

ANOVA Table for Two-Stage Nested Design

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
A	SS_A	$a-1$	$MS_A = SS_A/(a-1)$	$F_A =$ (see ‡ below)
$B(A)$	$SS_{B(A)}$	$a(b-1)$	$MS_B = SS_{B(A)}/[a(b-1)]$	$F_B = MS_{B(A)}/MS_E$
Error	SS_E	$ab(n-1)$	$MS_E = SS_E/[ab(n-1)]$	——
Total	SS_{total}	$abn-1$	——	——

‡ If $B(A)$ is a fixed factor then $F_A = MS_A/MS_E$
If $B(A)$ is a random factor then $F_A = MS_A/MS_{B(A)}$

1.2 ANOVA of nested design

Linear additive model

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

Where:

- Y_{ijk} = Observation
- μ = Overall mean
- α_i = effects of i^{th} level of Factor A (i = 1…a)
- β_{ij} = effects of j^{th} level of Factor B nested within i^{th} level of factor A (j = 1…b)
- ε_{ijk} = random error associated with k^{th} observation in treatment combination of j^{th} level of B nested within i^{th} level of A (k = 1…n)

1.2 ANOVA of nested design

Error term for factor A with fixed or random nested factor (factor B)

- The error term for big factor is different when the nested factor is random or fixed → depends on the expected mean square when the nested factor (Factor B(A)) is random or fixed.

nested factor is the small factor

Sources	Mean square (MS)	Nested factor (B(A)) is random		Nested factor (B(A)) is fixed	
		Expected mean square (EMS)	F test	Expected mean square (EMS)	F test
Factor A	MSA	$\sigma^2 + n\sigma_b^2 + \frac{bn\sum a_i^2}{a-1}$	$\frac{MSA}{MSB(A)}$	$\sigma^2 + \frac{bn\sum a_i^2}{a-1}$	$\frac{MSA}{MSE}$
Factor B(A)	MSB(A)	$\sigma^2 + n\sigma_b^2$	$\frac{MSB(A)}{MSE}$	$\sigma^2 + \frac{n\sum\sum b_{j(i)}^2}{a(b-1)}$	$\frac{MSB(A)}{MSE}$
Error	MSE	σ^2		σ^2	

Note: B(A) = B nested in A

2.0 Nested design with fixed and random effect

Example 1: Two factors - School and teacher

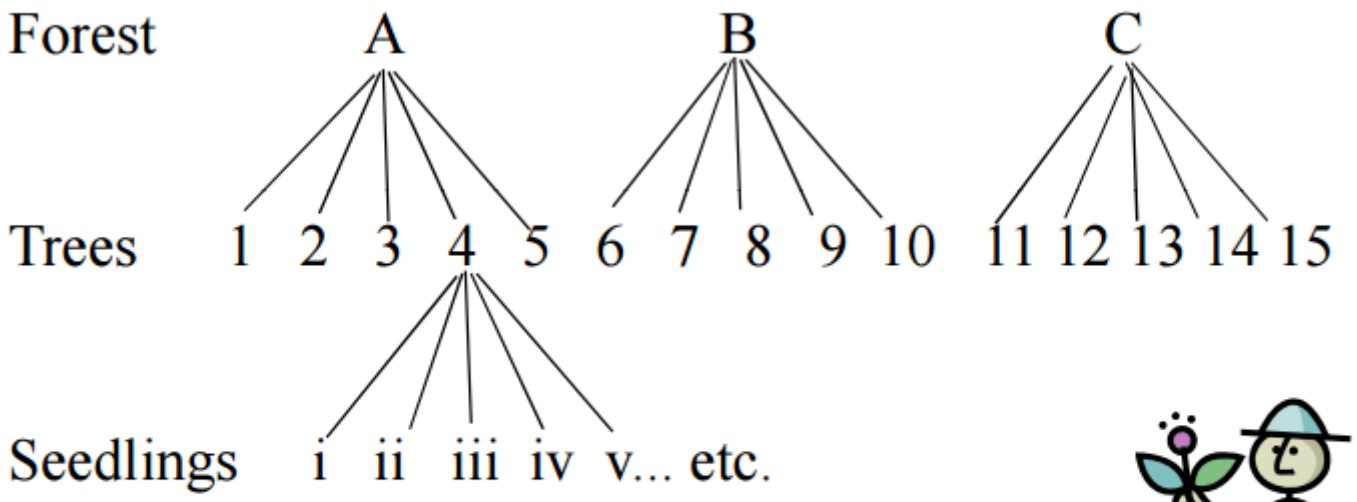
Objective:

To study the performance of school based on teachers who attended a course focusing on a new teaching method.

- School: fixed effect
- Teacher : random effect
  - There are 2 schools, A and B. Each school has 3 teachers under evaluation, teacher 1, 2 and 3.
  - In this case, nested means, that teacher 1 in school A is not the same as teacher 1 in school B, and so on.
  - This has to be kept in mind when trying to determine if the design is crossed or nested. To be crossed, the same teacher needs to teach at all the schools.

2.0 Nested design with fixed and random effect

Example 2: Two factors - Forest and seedling



Total of 75 seedlings.

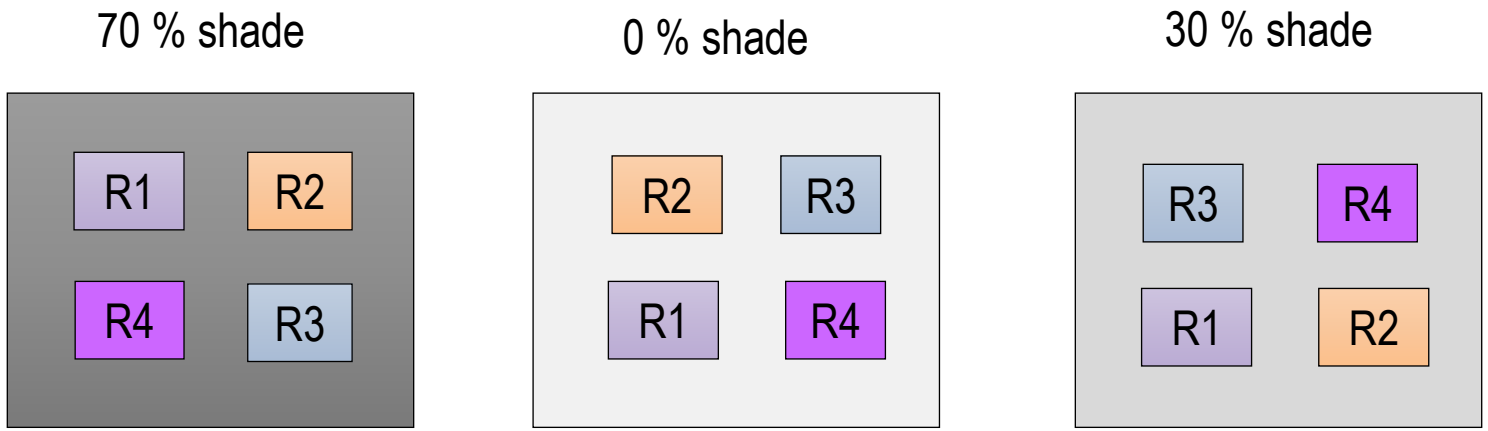
2.1 One-factor nested design – fixed model

what is fixed model?

Characteristics:
Both big and small factors are treated as fixed effects.
Interested in the mean for each factor level

Example: Shade percentage

- **Experimental setup:** The structure of treatment could not be replicated and randomized
- **Objective:** To study the effect of shading on internode length of turfgrass.
- For shade treatment, all experimental units were randomly placed under each % of shade.
- This experiment was conducted in a nested design.
- **Layout:**



2.1 One-factor nested design – fixed model

ANOVA table

The ANOVA table for one-factor nested design with fixed effects:

Source	df	SS	MS	F
Treatment	t -1		MStrt	MStrt/ MSE
Error	(N-1) – (t-1)		MSE	
Total	N – 1			

2.1 One-factor nested design – fixed model

Example dataset

Replication	0%	30 %	70%
1	3.55	5.98	10.45
2	4.03	5.50	10.12
3	3.15	5.06	9.96
4	3.66	5.59	10.38

2.1 One-factor nested design – fixed model

fixed model

R codes

```
setwd("D:/… 1/R")
nest_one <- read.csv ("nested.csv", sep = "," , header = TRUE )
str(nest_one)
as.factor(nest_one$source)

#fit model
fit_fixed <- lm(yld ~ shade, data = nest_one)
anova(fit_fixed)
#mean comparison
library (“agricolae”)
lsd_shade<- LSD.test(fit_fixed, “shade”)
```

2.2 Two-factor nested design – fixed model

Example: Nutrient formulation and planting media

- Objective: To study the effect of formulation and planting media on yield of tomato.
- Experimental set up:
  - This experiment is conducted under a greenhouse condition
  - The nutrient formulation is supplied through irrigation system from one big tank to all media (that means there is no replication made on the formulation).
- Thus, the experiment was conducted in a nested design.

two factors, formulation and media
note, formulation is not replicated, thus considered as random.

A nested design (hierarchical design) is used for experiments when the experimental units are sub-sampled.

And each subunit sample is unique!

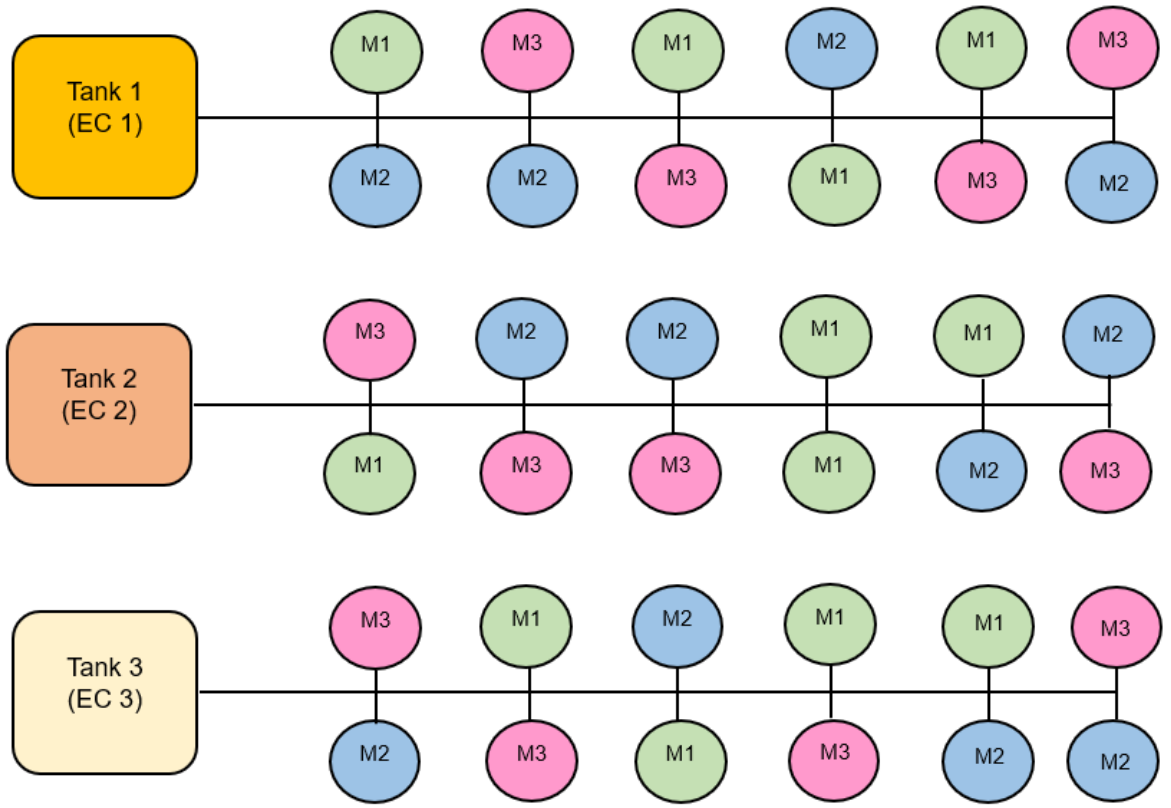
2.2 Two-factor nested design – fixed model

Design layout

- **Factor A**: Electrical conductivity (EC) of nutrient solution
  - Level: (EC1, EC2, EC3)
- **Factor B**: Media
  - Level: (M1, M2, M3)

- Note that **EC is not replicated**, only media is replicated
- In this experiment, media is nested within EC

EC (big factor) is random. Media (small factor) is fixed.



2.2 Two-factor nested design – fixed model

ANOVA table

Factor B is nested within A

Source	df	SS	MS	F
Factor A	a -1		MSA	MSA/ MSE
Factor B(A)	a(b -1)		MS A(B)	MSA(B)/ MSE
Error	ab (r-1)		MSE	
Total	abr – 1			

Notes:
r = number of replication
abr = N (total observation)

2.2 Two-factor nested design – fixed model

Example data

- Objective: To study the effect of formulation and planting media on yield of tomato.

Rep	Electrical conductivity (EC)								
	EC 1			EC 2			EC 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
1	1.25	1.44	1.46	2.06	3.02	5.98	6.96	6.25	5.70
2	1.55	1.47	1.46	2.98	3.40	5.42	7.45	6.33	5.45
3	1.66	1.33	1.41	2.59	3.55	5.51	7.38	6.12	5.66
4	1.03	1.36	1.27	2.50	2.98	5.12	7.12	6.13	6.16

2.2 Two-factor nested design – fixed model

R codes

```
setwd("D:/… 1/R")
nest_two <- read.csv ("nested.csv", sep = "," , header = TRUE )
str(nest_two)
nest_two$EC <- as.factor(nest_two$EC)
nest_two$media <- as.factor(nest_two$media)

fit_two <- lm(y ~ EC + EC/media, data = nest_two)
anova(fit_two)

#mean comparison
library (“agricolae”)
lsd_EC<- LSD.test(fit_two, “EC”)
```

2.3 Two-factor nested design – mixed model

Example: Shade and varieties

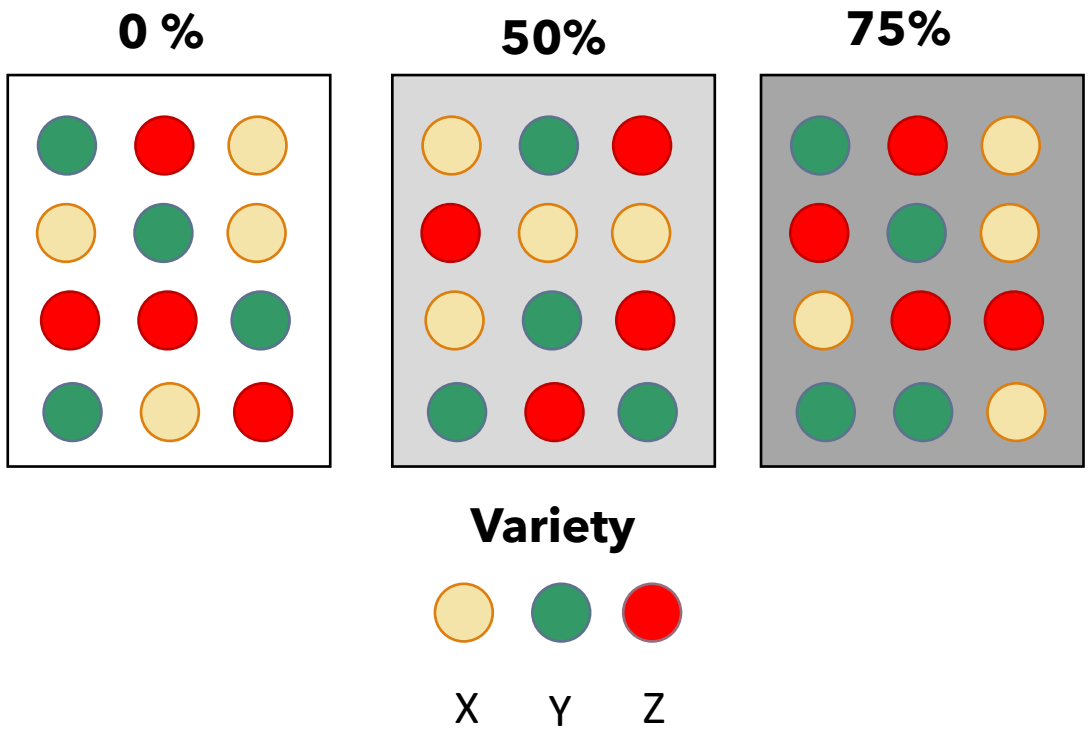
- **Mixed model** →A is fixed, and B(A) is random

- **Two factors:**
  - Big factor → Shades: 0, 50%, 75% - Fixed Effects
  - Small factor → Varieties: X, Y, Z – Random Effects
- Replication: 4 reps for each variety within shade

Note: In this experiment, the researcher is not interested to know the effects of each variety. The variety is only as representative in the experiment. Thus, variety is treated as random effect.

If interested to know the effects of variety, treat variety as fixed, then use fixed model ANOVA.

Design Layout



2.3 Two-factor nested design – mixed model

ANOVA table – nested factor is random

Shade: 3
Varieties: 3
Rep: 4

DF for variety(shade) = (3-1) variety x (3) shade

Source	df	SS	MS	F
Shade (S)	s -1 3-1 =2		MSShade	MSShade/ MSVar(Shade)
Var (Shade)	s(v -1) 3 (3-1) =6		MSVar (Shade)	MSVar (Shade)/ MSError
Error	sv(r-1) 3*3(4-1) = 27		MSError	
Total	vsr – 1 (3*3*4) -1= 35			

Var (Shade) = Variety within shade

2.3 Two-factor nested design – mixed model

R codes

```
setwd("D:/… 1/R")
nested <- read.csv ("nested.csv", sep = "," , header = TRUE)
str(nested)
nested$shade <- as.factor(nested$shade)
nested$variety <- as.factor(nested$variety)

#nested factor are random
fit_mixed<- aov(y ~ shade + Error(variety), data = nested))
Summary (fit_mixed)
```

Specify the error term for the random nested factor

The dataset format of nested design when nested factor is random

The level for random nested factor B(A) – variety, should continue from one level of factor A to another level as shown below:

Shade : 0%, 50% and 75% (big factor)

Variety (nested factor):
- (v1,v2,v3) – shade 0%
- (v4, v5, v6) – shade 50%
- (v7, v8, v9) – shade 75%

Thus, there are
- 3 levels of factor A
- 9 levels of factor B(A)

	A	B	C
1	shade	var	y
2	0	v1	15.34
3	0	v1	15.02
4	0	v1	15.01
5	0	v1	15.93
6	0	v2	16.73
7	0	v2	16.26
8	0	v2	17.02
9	0	v2	16.64
10	0	v3	15.65
11	0	v3	15.21
12	0	v3	15.58
13	0	v3	15.56
14	50	v4	16.88
15	50	v4	17.28
16	50	v4	17.07
17	50	v4	16.96
18	50	v5	17.38
19	50	v5	18.25
20	50	v5	17.93
21	50	v5	17.82
22	50	v6	16.47
23	50	v6	17.06
24	50	v6	16.64
25	50	v6	16.47
26	75	v7	25.35
27	75	v7	25.59
28	75	v7	25.66
29	75	v7	25.40
30	75	v8	26.61
31	75	v8	26.43
32	75	v8	27.15
33	75	v8	26.73
34	75	v9	25.17
35	75	v9	25.27
36	75	v9	25.72
37	75	v9	25.98

3.0 Difference between Nested & Split Plot

Question:

What is the difference between split plot and two-factor nested design?

- Layout
- ANOVA

3.0 Difference between Nested & Split Plot - ANOVA

NESTED

- Big factor=A; small factor=B
- No blocks. Only nested factor is replicated
- Interaction effects cannot be measured

Source	df	SS	MS	F (B Fixed)	F (B random)
Factor A	a -1		MSA	MSA/ MSE	MSA / MSB(A)
Factor B(A)	a(b -1)		MSB (A)	MSA(B) / MSE	
Error	ab (r-1)		MSE		
Total	abr – 1				

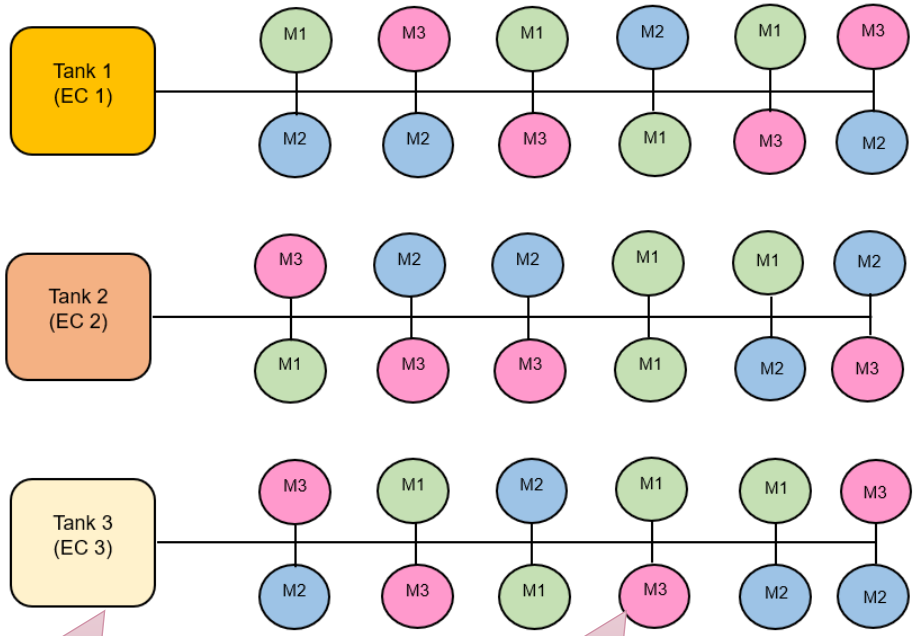
SPLIT PLOT

- Main Plot=A; Sub Plot=B
- Must have BLOCKS – for RCBD
- Error term for Main Plot = Block x A
- Interaction effects can be measured

Source	Degree of freedom	Sum of square	Mean square	F test
Block	r-1	SSR	MSR	MSR/$\mathbf{MSE_A}$
Factor A (Main plot)	a-1	SSA	MSA	MSA/$\mathbf{MSE_A}$
Error a (Block* A)	**(r-1)*(a-1)**	**SSE_A**	**MSE_A**	
Factor B	b-1	SSB	MSB	MSB/$\mathbf{MSE_B}$
Factor A*Factor B	(a-1)*(b-1)	SSAB	MSAB	MSAB/$\mathbf{MSE_B}$
Error b	**a(r-1)(b-1)**	**SSE_B**	**MSE_B**	
Total	(rab) -1	SST		

3.0 Difference between Nested & Split Plot - Layout

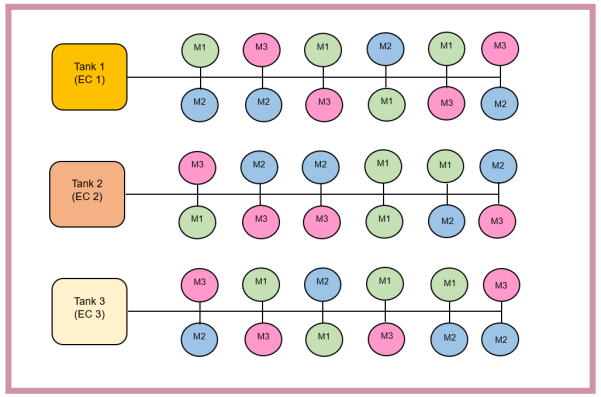
Nested design (2-factor)



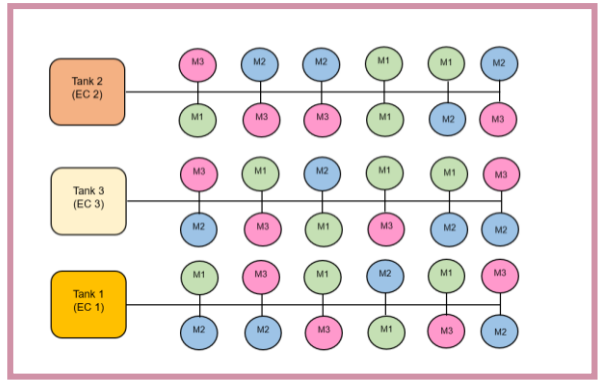
Factor 1 is not replicated

Factor 2 is replicated

Split plot design (RCBD)



Rep 1 (Block 1)



Rep 2 (Block 2)

Both factors are replicated