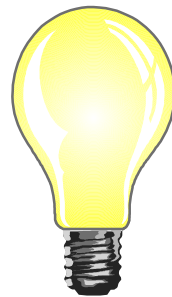


Statistical Techniques II

EXST7015

Experimental Design



Experimental Design

- **The design aspect of Analysis of Variance refers to fate of the error term.**
- **We will discuss three designs,**
 - ▶ **Completely Randomized Design**
 - ▶ **Randomized Block Design**
 - ▶ **Latin Square Design**
- **We will also discuss hierarchical design, or nested error terms. Any of the three designs above can have nested error terms.**

Completely Randomized Design (CRD)

- The basic, most simple design is the CRD (Completely Randomized Design) with a single error term.

- ▶ $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

- We will talk about the nature and arrangement of treatments later. First we talk about the error term. Differences among designs come primarily from the error terms.

Schematic of a CRD

A	B	A	C	B
C	B	B	A	C
B	C	A	C	A

CRD (*continued*)

- Depending on how the experiment was done we may find that the error term is not a simple randomized, replicated observation with a treatment.
- There may be several other sources of variation within the error term, which if ignored will alter the error term and reduce the effectiveness of the experiment.

CRD (*continued*)

- The source table for this CRD, with sources, degrees of freedom and EMS is given below. The treatments can be fixed or random, but the error term must be a random effect.

Source	d.f.	SS	MS	EMS
Tmt	t-1	SSTmt	MSTmt	$\sigma^2 + n\sigma^2_{\tau}$
Error	t(n-1)	SSE	MSE	σ^2
Total	tn-1	SSTotal		

CRD (*continued*)

- CRD with fixed effect treatments.
- Fixed effects include all levels of interest for the treatment, or all possible levels of the treatment.

Source	d.f.	SS	MS	EMS
Tmt	t-1	SSTmt	MSTmt	$\sigma^2 + \frac{n \sum \tau_i^2}{(t-1)}$
Error	t(n-1)	SSE	MSE	σ^2
Total	tn-1	SSTotal		

CRD (*continued*)

- The EMS translate directly into the F tests of treatments.

- Random effects

$$F = \frac{\sigma^2 + n\sigma_{\tau}^2}{\sigma^2}$$

- Fixed effects

$$F = \frac{\sigma^2 + n\sum \tau_i^2 / t - 1}{\sigma^2}$$

Fixed and Random effects

- **In more advanced designs which treatments are random and which are fixed becomes extremely important.**
- **However, for the moment we will only be putting one treatment in our models, and it makes little difference if that treatment is fixed or random.**
- **We will continue to determine if each treatment is fixed or random because of its eventual importance.**

Sampling Units

- In a CRD the treatments are assigned, completely at random, to some unit. That unit may be a plot in a field, a test tube, a plant in a pot, a car, a laboratory rat, a person, just about anything.
- The unit that we assign the treatment to is called the **experimental unit**. The error term derived from the variance of that unit is called the experimental error.

Sampling Units (*continued*)

- The actual measurement of the dependent variable Y_{ij} , is done on the experimental unit, or some smaller unit than the experimental unit.
- This smaller unit is called the **sampling unit** or measurement unit.
- For example, the treatment is applied to a plot in a field. This is the experimental unit. We may measure the height of one plant out of many in that plot, this is the sampling unit.

Sampling Units (*continued*)

- **In practice, if we only have one measurement per experimental unit we consider this to represent the experimental unit, even if we measure some smaller unit.**
- **In this case the experimental unit and sampling unit are the same unit.**
- **If we measure only one plant in the plot, that plant represents the whole plot. If we take only one blood sample from a rabbit, it represents the whole rabbit.**

Sampling Units (*continued*)

- **The importance of sampling units only comes into play when we have several measurements of sampling units per experimental unit.**
- **These are nested error terms.**

Nested error terms (*continued*)

- For example, suppose we have 4 fertilizer treatments ($t=4$) in a field.
- Each treatment level occurs in 5 randomly selected plots ($p=5$), so the field has 20 plots. **This is the experimental unit.**
- We want to measure Plant available phosphorus. The design is a simple CRD if there is only one measurement per experimental unit.

Nested error terms (*continued*)

- **But suppose we proceed to take 3 soil samples ($s=3$) in each plot, and back in the laboratory we make 2 measurements on each soil sample ($n=2$).**
 - ▶ **Treatment : fertilizer ($t=4$)**
 - ▶ **Experimental unit : a plot ($p=5$)**
 - ▶ **Sampling unit : a soil sample ($s=3$)**
 - ▶ **Sub-sampling unit : a soil test ($n=2$)**

Nested error terms (*continued*)

- This is hierarchical because we took the soil samples as sampling units of the plots, and then made several measurements that are sub-sampling units of each soil sample.
- The model is
 - ▶ $Y_{ijkl} = \mu + \tau_i + \delta_{ij} + \gamma_{ijk} + \varepsilon_{ijkl}$
- Notice that the subscripts on the terms will show the hierarchical nature of the sampling.

Nested error terms (continued)

- Expected mean square structure.
- For fixed treatments, replace σ^2_τ with $\Sigma\tau^2_i/(t-1)$.

Source	d.f.	EMS
Tmt	t-1	$\sigma^2 + n\sigma^2_\gamma + ns\sigma^2_\delta + nsp\sigma^2_\tau$
Plot(Tmt)	t(p-1)	$\sigma^2 + n\sigma^2_\gamma + ns\sigma^2_\delta$
Sample (Plot*Tmt)	tp(s-1)	$\sigma^2 + n\sigma^2_\gamma$
Measure (sample*plot*tmt)	tps(n-1)	σ^2
Total	tpsn-1	

Nested error terms (continued)

- **This model has a hierarchical or nested expected mean square structure.**
 - ▶ **The Plot(Tmt) is the experimental error.**
 - ▶ **The Sample(plot tmt) is the sampling error.**
 - ▶ **The measure(sample plot tmt) is the sub-sampling error.**
- **The last error is also called the residual error. In SAS it will contain any terms left off the model.**

Nested error terms (*continued*)

- Note that the appropriate error term for testing treatments is the Plot(Tmt) error term.
- We expect, under the null hypothesis, that the numerator and denominator of an F test are the same.
- If we want to test $H_0: \sigma^2_{\tau} = 0$ in a term consisting of $\sigma^2 + n\sigma^2_{\gamma} + ns\sigma^2_{\delta} + nsp\sigma^2_{\tau}$, we need an error term containing $\sigma^2 + n\sigma^2_{\gamma} + ns\sigma^2_{\delta}$.

Nested error terms (*continued*)

- We want to test the term σ^2_τ , ($H_0:\sigma^2_\tau=0$), so we must find an error term that has all the same terms as the treatments except this one.
- Power is gained because the coefficient of σ^2_τ is "nsp" instead of just "n".

$$F = \frac{\sigma^2 + n\sigma_\gamma^2 + ns\sigma_\delta^2 + nsp\sigma_\tau^2}{\sigma^2 + n\sigma_\gamma^2 + ns\sigma_\delta^2}$$

Nested error terms (*continued*)

- **We started with treatments and experimental units. We added sampling and sub-sampling units.**
- **We find that we must do our test with the experimental error.**
- **So what did we gain? We would have used this anyway, even if we had not taken all the other measurements.**
- **We gained power because of the increased coefficient "nsp" on the σ^2_{τ} term.**

Randomized Block Design

- **Blocks:** In some experiments the whole experiment cannot be conducted in a single place or at a given time.
 - ▶ We may find that we have to use 3 incubators for our cultures because they don't fit in one incubator.
 - ▶ We may have to use several different fields to conduct the experiment if a single large field is not available.
 - ▶ We may have to repeat our experiment several times if we cannot do it all at once.

RBD (*continued*)

- **These incubators or fields or times are not the same. They differ in some way.**
- **If we ignore this variation, the extra variation will inflate our error term.**
- **How do we get this extra variation out of the error? We put it in the model. We will call this a BLOCK.**

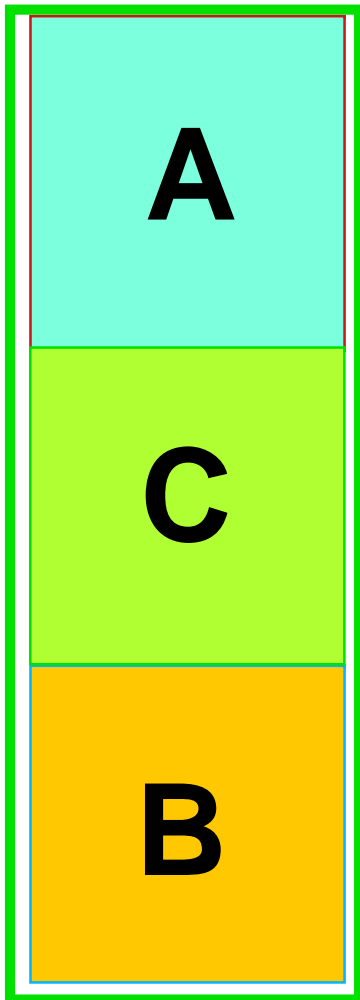
RBD (*continued*)

- A block is NOT a source of variation that we are interested in interpreting. We simply recognize that it exists and include it in the model to remove it from the error.
- It will be put in the model and appear just like a treatment.
- The model would be

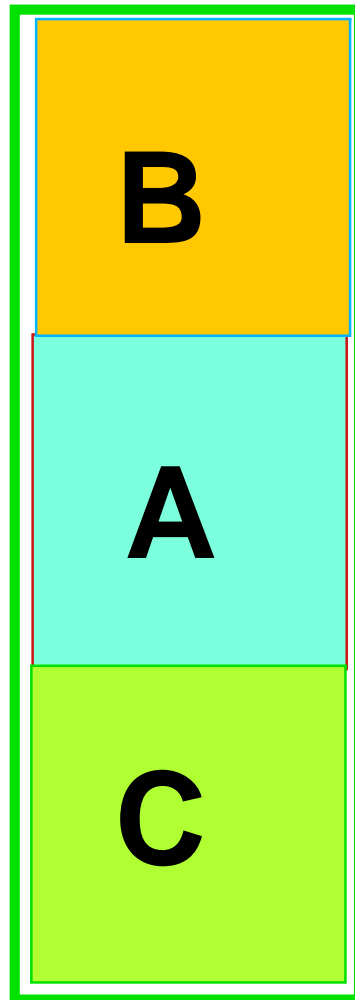
$$\blacktriangleright Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

Schematic of an RBD

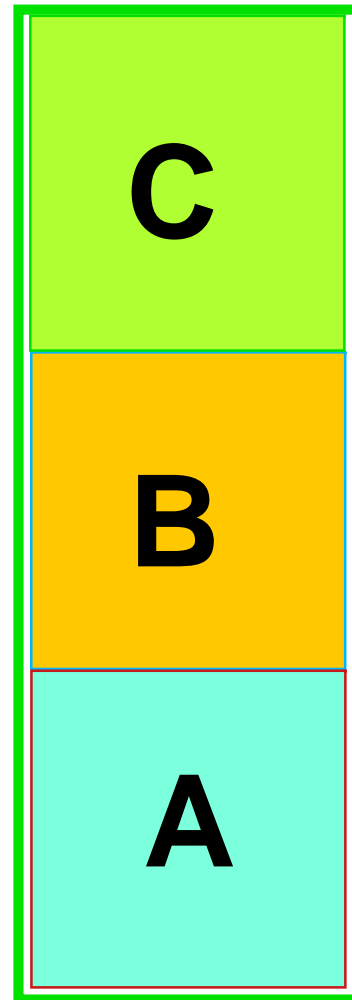
Block 1



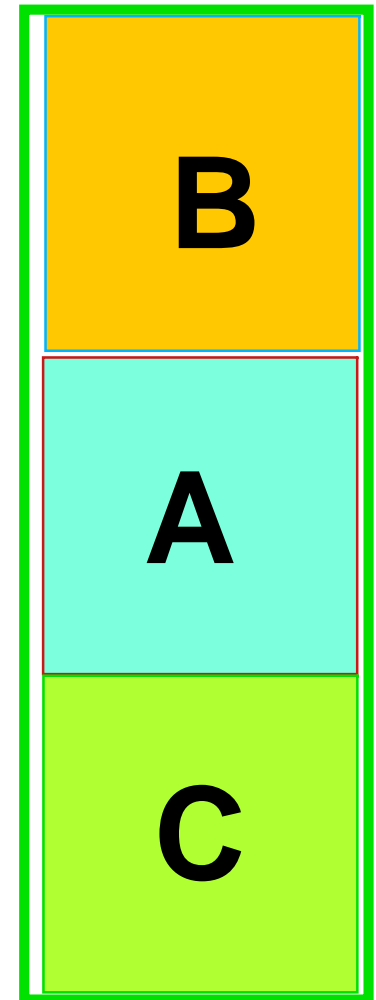
Block 2



Block 3



Block 4



RBD (*continued*)

- The source table for this CRD, with sources, degrees of freedom and EMS is given below.

Source	d.f.	SS	MS	EMS
Tmt	t-1	SSTmt	MSTmt	$\sigma^2 + n\sigma^2_{\tau}$
Block	b-1	SSBIk	MSBIk	$\sigma^2 + n\sigma^2_{\beta}$
Error	(t-1) (b-1)	SSE	MSE	σ^2
Total	tn-1	SSTotal		

RBD (*continued*)

- **The block design depicted has only one cell for each treatment per block. This is a valid and legitimate design, since the "interaction" of blocks and treatments serves as an error term.**
- **If the behavior of the treatments between blocks is consistent, this error term is a measure of the variability of experimental units between blocks.**
- **Why does it serve as an error term?**

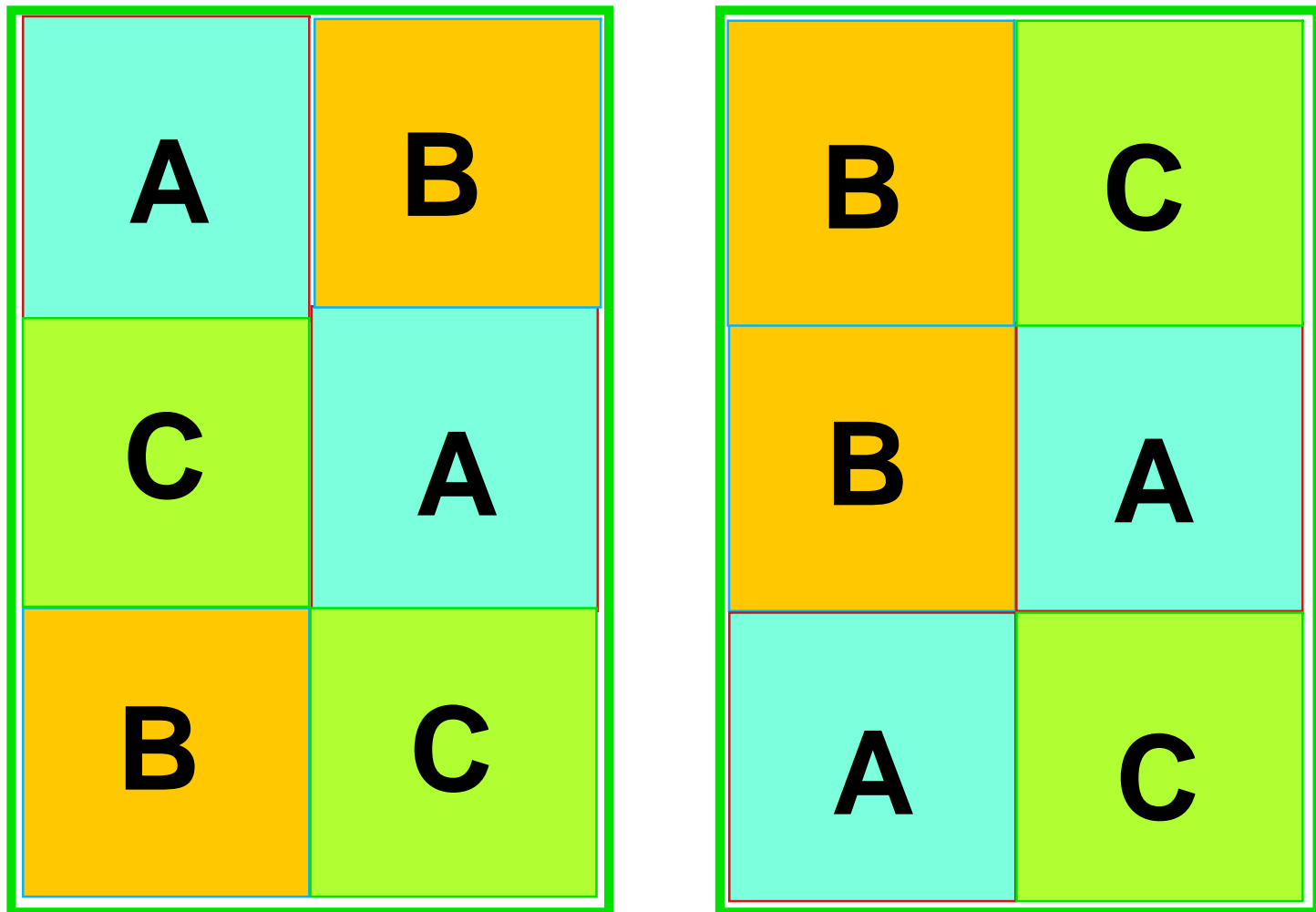
RBD (*continued*)

- **The interaction serves as an error term because it is a random effect. It is a random effect because the block is random (almost invariably), and interactions of random effects are random.**
- **Treatments can be fixed or random, but as long as one term in an interaction (the block) is random, the interaction is random.**

RBD (*continued*)

- Note that the degrees of freedom for an interaction between treatments (t) and blocks (b) is $(t-1)(b-1)$. This is typical of interactions in general.
- How would the design differ if we had replicated experimental in the blocks?
- These would be nested.

Schematic of an RBD with replication within the cells



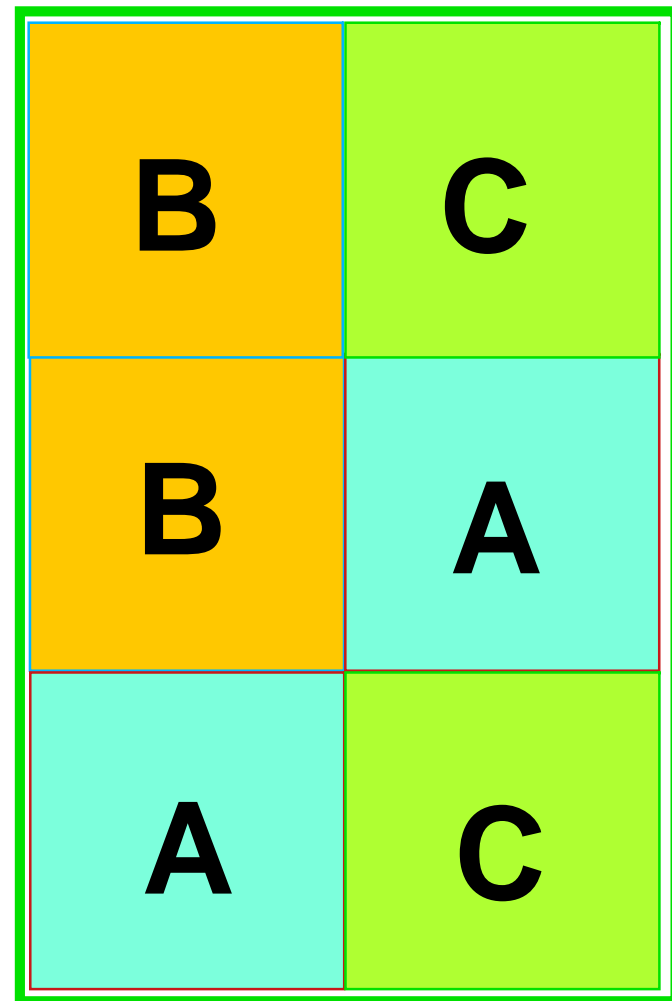
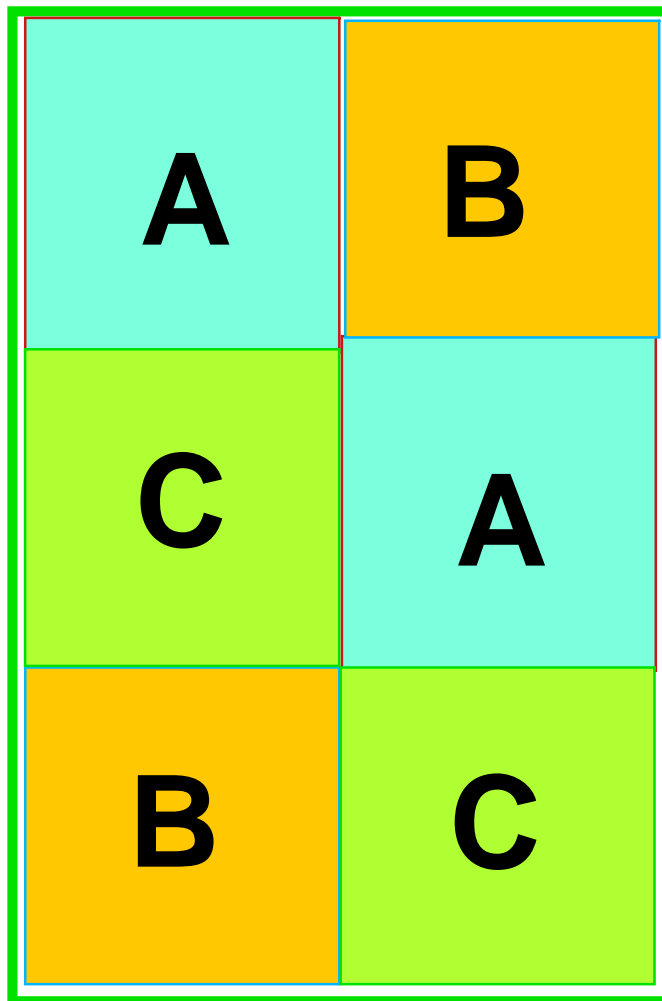
RBD (*continued*)

- For this design we still have treatments, blocks and an interaction. However, we now also have a measure of variability of treatments within blocks as well as between blocks.
- The model would be
 - ▶ $Y_{ij} = \mu + \beta_i + \tau_j + \beta\tau_{ij} + \varepsilon_{ijk}$

RBD (*continued*)

- **There are actually two ways this model can arise, and they are different in terms of what can be tested.**
- **In the case just discussed, the plots were the experimental units. The "interaction" between treatments and blocks represent variability among experimental units.**
- **The replicated plots in each block also represent variability among experimental units.**

Replicated cells in blocks

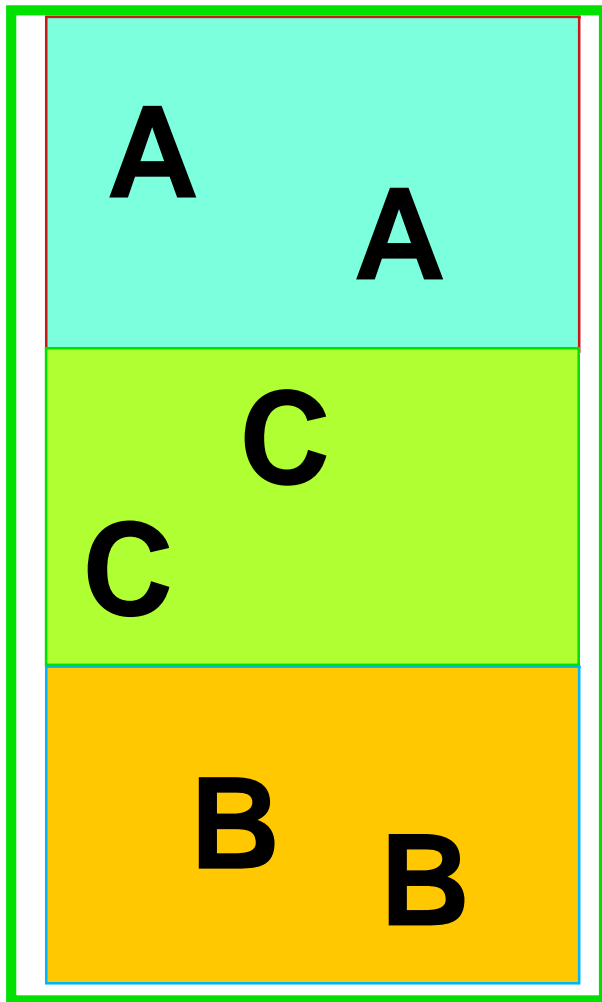


RBD (*continued*)

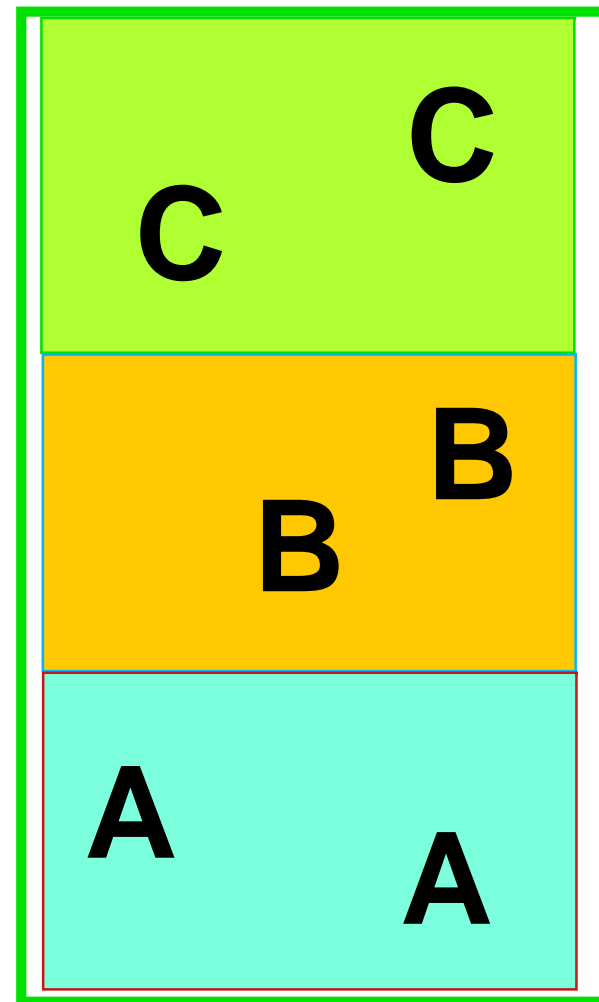
- **If we have two measurements of experimental units, they should be equal and we can test to see if they are equal.**
- **Another possibility is that the replicated measurements come from sampling units within the plots.**

Replicated samples in cells

Block 1



Block 2



RBD (*continued*)

- In this case we do not expect the sampling units from within plots to estimate the same variability as the between experimental unit variation.
- The plots estimate experimental error.
- The replicate measurements estimated sampling error.
- Though the model may look the same.

$$\blacktriangleright Y_{ij} = \mu + \beta_i + \tau_j + \beta\tau_{ij} + \varepsilon_{ijk}$$

RBD (*continued*)

Source	d.f.	SS	EMS
Tmt	t-1	SSTmt	$\sigma^2 + n\sigma_{\tau\beta}^2 + nb\sigma_{\tau}^2$
Block	b-1	SSBlk	$\sigma^2 + n\sigma_{\tau\beta}^2 + nt\sigma_{\beta}^2$
Tmt*Blk	(t-1)(b-1)	SSB*T	$\sigma^2 + n\sigma_{\tau\beta}^2$
Error	tb(n-1)	SSE	σ^2
Total	tbn-1	SSTotal	

RBD (*continued*)

- **Note that in this case the two error terms were both supposed to represent experimental error if the second error was replicated plots.**
- **Or the first represented experimental error and the second sampling error if the replicated measurements were taken from within plots.**
- **In either case there is a test of hypothesis here, but with different interpretations.**

RBD (*continued*)

- **If the two terms both represent experimental error, then we consider the possibility that for some reason the treatments do not behave the same way in the different blocks.**
- **In this case the " $\tau\beta$ " interaction represents a true interaction. Hopefully this does not exist, but since we have two estimates of experimental error we can test this ($H_0: \sigma^2_{\tau\beta} = 0$).**

RBD (*continued*)

- In the other case the two terms represent an experimental error and a sampling error. We expect that the error from within the more homogeneous plots would be smaller than between the plots.
- In this case the sampling error is σ^2 , and the experimental error is represented by the $\sigma^2 + n\sigma_{\tau\beta}^2$ term.
- We can test to see if these are the same, but we do not really expect them to be the same.

RBD (*continued*)

- In this last case we cannot test to see if there is a true interaction between the blocks and the treatments.
- In this case we can only **ASSUME** that there is not block by treatment interaction, and in fact this is a **new assumption for block designs.**

RBD (*continued*)

- **In the first case, where the two error terms both represented experimental error, we may consider the possibility that the two error terms be combined into a single error term.**
- **Is this wise? More d.f., more power. But are the error terms really the same? And which one is better if they are not the same?**

RBD (*continued*)

- **If the two are not the same, then the difference is caused by an "interaction: between the block and treatment.**
- **This means that for some reason the treatments did not give a consistent performance in the various blocks.**
- **Lets suppose we had an experiment to find the better of 5 rice varieties (A, B, C, D and E). We did the experiment each year for 4 years, and will block on years.**

RBD (*continued*)

- **If there is an interaction between blocks and treatments, it implies that some rice varieties did relatively better in some years and other varieties did better in other years!**
- **Suppose variety A did better in 1994 because it was a dry year, and variety B did better in 1995 (a wet year). Which one are you going to recommend to a farmer in 2001?**

RBD (*continued*)

- You don't know, because you don't know if 2001 will be wet or dry, or in between.
- So if you are going to conclude that one variety is "better", it should be a difference that is consistent across years, or a difference that considers the annual variation and interaction. This would be the $\sigma^2 + n\sigma^2_{\tau\beta}$ term.
- This will be a larger term, and harder to show a difference with, but the difference will be more certain.

RBD (*continued*)

- **On the other hand, if the errors are the same, why not pool?**
- **We need a mechanism to determine if we should pool or not.**
- **Obviously if the treatments are significantly different, we use the interaction term and do not pool.**

RBD (*continued*)

- **But if the two terms are not significantly different, are they the same?**
- **We cannot show statistically that two things are the same because we do not know the probability of Type II error.**
- **So how similar do they have to be before we would pool?**
- **See pooling criteria by Bancroft and Chien-Pai (JASA, 1983, 78(384):981-983).**

RBD (*continued*)

Values are $P(>F)$. Pool if equal to or larger than the values in the table.

	n1 = 4	8	12	16	20
n2 = 4	0.35	0.43	0.45	0.48	0.48
8	0.29	0.37	0.40	0.43	0.43
12	0.26	0.34	0.37	0.40	0.41
16	0.25	0.32	0.36	0.38	0.39
20	0.24	0.31	0.34	0.38	0.38

RBD (*continued*)

- **So if we have two estimates of experimental error (between plots), we may wish to pool.**
- **If we have one estimate of experimental error and a sampling error (within plots), we are less likely to want to pool.**
- **Later we will see that when we have several treatment terms, each with a block interaction, we will usually pool all block interactions into a single error term.**

RBD (*continued*)

- **Could we have more than one block?**
- **Sure, if we have several fields, each sampled in several years, we could block on both years and fields.**
- **Note that for t treatments in y years and f fields, we should have $t*y*f$ experimental units. For example, with 5 treatments in 4 fields over 3 years we should have $5*4*3=60$ experimental units. This is a minimum. Replicated experimental units would be extra.**

Latin Square Designs (LSD)

- **Suppose we had 5 diet treatments that we wanted to examine for their influence on milk production. Our department provided us with 5 cows to do our experiment.**
- **Five treatments, 5 cows, no reps?**
- **We could block on time, and do the experiment over several weeks with weekly estimates of total milk production. That might work.**

LSD (*continued*)

- **However, there is a little problem. The cows are different. They have different milk production rates. They always have had and always will have. We could try to look at pre-post results, the change in milk production from before the diet to after the diet. That might work.**
- **But there is another way.**
- **The LATIN SQUARE DESIGN.**

LSD (*continued*)

- In the Latin Square Design each cow will get each diet, so cow differences average out, and won't affect the results.
- Obviously, with 5 diets and 5 cows we have to do the experiment for 5 weeks to give each cow each diet.
- A Latin Square has a special setup so that each diet occurs in each week (weeks may differ) and with each cow (to average out cow differences).

LSD (*continued*)

The Latin Square below has not been randomized. To randomize the diet rows would be placed in random order and then the diet columns would be placed in random order.

	Cow1	Cow2	Cow3	Cow4	Cow5
Week 1	A	B	C	D	E
Week 2	B	C	D	E	A
Week 3	C	D	E	A	B
Week 4	D	E	A	B	C
Week 5	E	A	B	C	D

LSD (*continued*)

- Note that we have 5 diets, 5 cows and 5 weeks. The usual block design would require $5*5*5=125$ experimental units.
- However, we only have 25. This will only work (well) if we have a Latin Square arrangement with each diet occurring once with each cow and once in each week.

LSD (*continued*)

- The Latin Square source table. Note that $r=c=t$ for any Latin Square.

Source	d.f.	SS	EMS
Row	$r-1$	SSRow	$\sigma^2 + r\sigma^2_\rho$
Col	$c-1$	SSCol	$\sigma^2 + r\sigma^2_\beta$
Tmt	$t-1$	SSTmt	$\sigma^2 + r\sigma^2_\tau$
Error	$(r-1)(r-2)$	SSE	σ^2
Total	r^2-1	SSTotal	

LSD (*continued*)

- **The Latin square is a bit messy. We cannot examine any interactions at all because there are not enough degrees of freedom.**
- **Essentially the remains of any interactions are pooled into an error term. This is consistent with other design practices since ROW and COL are blocks and we usually pool block interactions into a single error term.**

LSD (*continued*)

- The error term for the Latin Square is usually best calculated as the remaining d.f. after subtracting the main effects from the total.
- The model is
 - ▶ $Y_{ij} = \mu + \beta_i + \rho_j + \tau_k + \varepsilon_{ij}$
- Note the odd subscripting. Each observation can be identified by just 2 subscripts.

Terminology

- **Some terminology**
- **Main effects are the lone treatments and blocks. The remaining terms are interaction terms or nested terms.**
- **Interactions are between two main effects that are cross classified.**

Terminology (*continued*)

- **Cross-classified effects are distinct, meaningful sources of variation and must kept in the appropriate categories.**

	Tmt 1	Tmt 2	Tmt 3	Tmt 4
Block 1	Y_{11}	Y_{12}	Y_{13}	Y_{14}
Block 2	Y_{21}	Y_{22}	Y_{23}	Y_{24}
Block 3	Y_{31}	Y_{32}	Y_{33}	Y_{34}
Block 4	Y_{41}	Y_{42}	Y_{43}	Y_{44}
Block 5	Y_{51}	Y_{52}	Y_{53}	Y_{54}

Terminology (*continued*)

- Nested effects are randomly applied (such as observations) and could be reordered without affecting the experiment.

	Tmt 1	Tmt 2	Tmt 3	Tmt 4
Obs 1	Y_{11}	Y_{12}	Y_{13}	Y_{14}
Obs 2	Y_{21}	Y_{22}	Y_{23}	Y_{24}
Obs 3	Y_{31}	Y_{32}	Y_{33}	Y_{34}
Obs 4	Y_{41}	Y_{42}	Y_{43}	Y_{44}
Obs 5	Y_{51}	Y_{52}	Y_{53}	Y_{54}

Series of Latin Squares

- **Latin squares are rather limited as to how they are done. Additional experimental units are not readily added.**
- **However, it is possible to add a second or third square, and reproduce the whole experiment elsewhere or at a different time.**

Series of LSD (*continued*)

■ Example

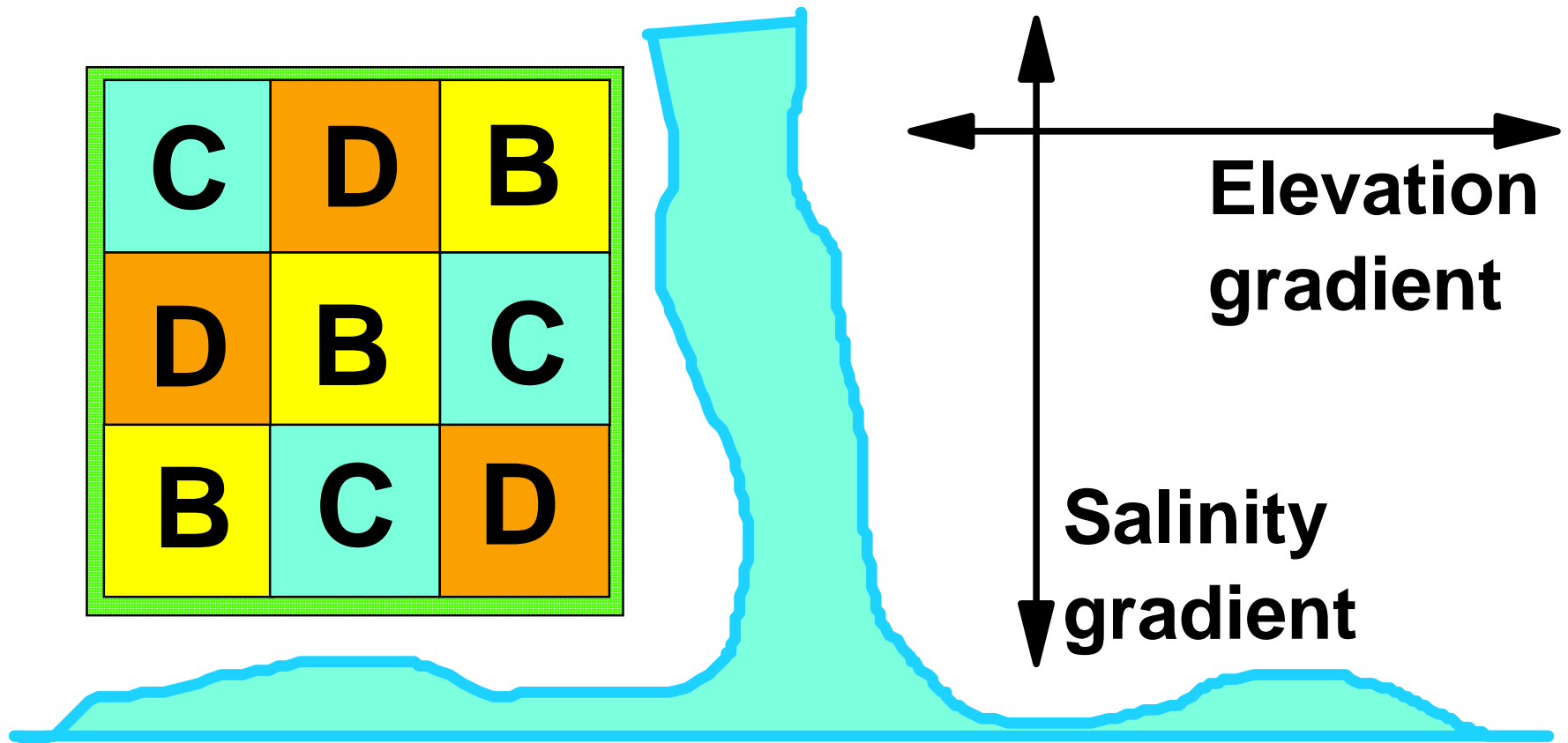
- ▶ **Suppose we are examining the effect of various treatments in removing oil from marsh area that have been fouled. We have 3 treatments we are comparing. The experimental area will be sprayed with oil and one of the following treatments applied.**
- ▶ **1) Control (no treatment), 2) detergent spray and 3) biological agent**

Series of LSD (*continued*)

- **The objective is to compare the effects of the treatments after two months. The variable to be measured is live *Spartina* biomass in the treatment plots.**
- **The marsh area to be used in the experiment has several gradients. There are saline gradients and elevation gradients that will affect the experiment. The investigators decided to "block" in both a North-South and East-West direction. This is a Latin Square.**

Series of LSD (*continued*)

- Layout of the experiment.

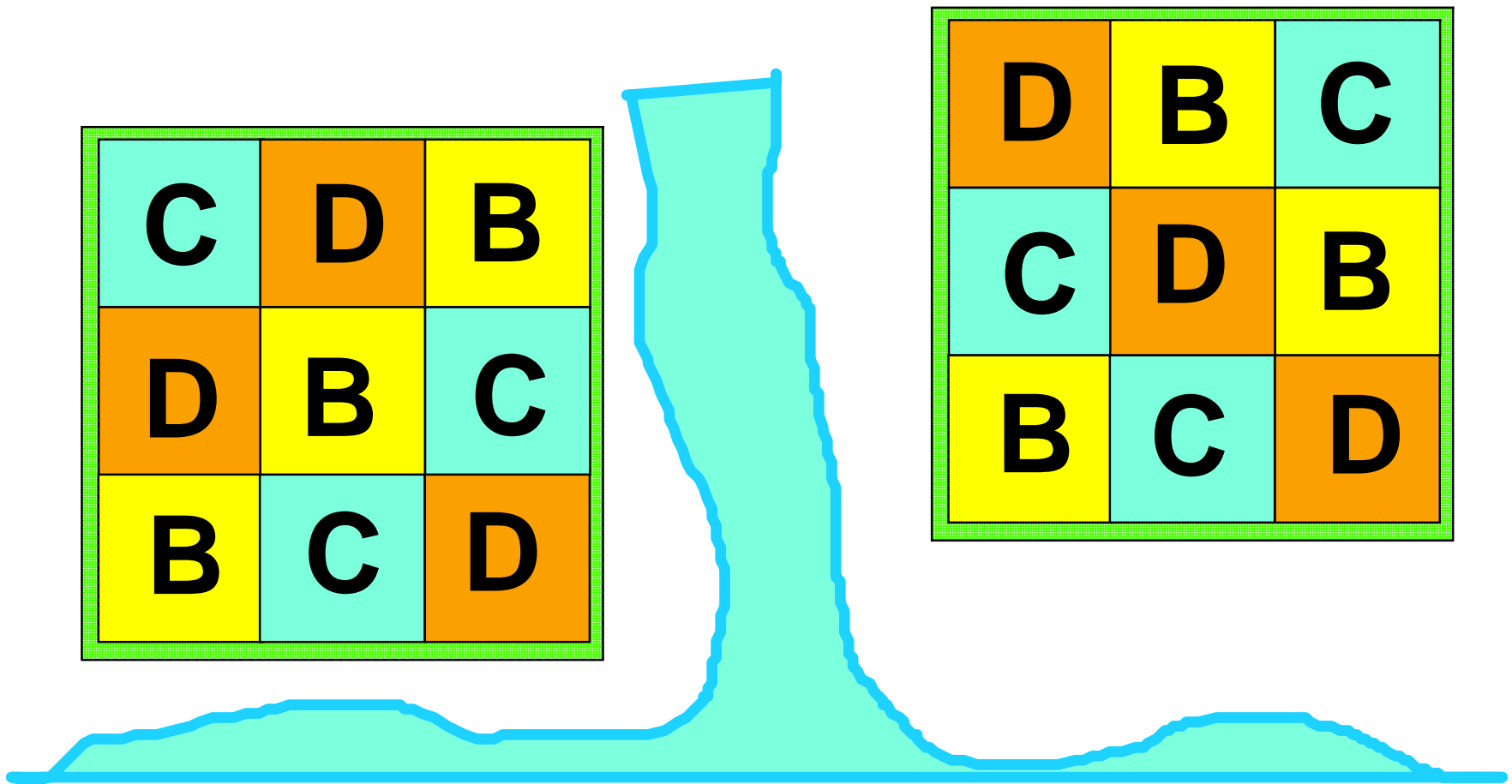


Series of LSD (*continued*)

- **There is nothing wrong with this experiment.**
- **Blocking on "rows" and "columns" should account for the salinity and elevation gradients.**
- **However, if the investigators decide they need additional replication, they could do another square elsewhere, perhaps across the stream.**

Series of LSD (*continued*)

- Layout of the expanded experiment.



Series of LSD (*continued*)

- **Now we need a source table. We still have the basic Latin square, but there are two squares. Call this variable "Square" with two levels, east and west.**
- **We are not interested in Square as a source of variation. It is simply a mechanism to increase replication, as blocking often is.**
- **However, it is a source of variation and must be included in the model.**

Series of LSD (*continued*)

- **Square definitely has meaning, east and west are two distinct squares. The treatments still have the same meaning in each square, so these are cross classified.**
- **Do rows and columns mean the same in the two squares? Does row 1 in the east have the same salinity as row 1 in the west. Does col 2 have the same salinity and elevation in both squares?**

Series of LSD (*continued*)

- **Probably not.** If row 1, 2 and 3 in the east has a different meaning from row 1, 2 and 3 in the west, these should be nested.
- **The same for columns.**
- **The model is**
 - ▶ $Y_{ijk} = \mu + \gamma_i + \beta_{ij} + \rho_{ik} + \tau_l + \gamma\tau_{il} + \varepsilon_{ijk}$

Series of LSD (*continued*)

- The source table (EMS later)

Source	d.f.	d.f. num	SS
Square	$s-1$	1	SSSquare
Row(Square)	$s(r-1)$	4	SSRow
Col(Square)	$s(c-1)$	4	SSCol
Tmt	$t-1$	2	SSTmt
Tmt*Square	$(t-1)(s-1)$	2	SSSquare*Tmt
Error	$s(r-1)(r-2)$	4	SSE
Total	sr^2-1	17	SSTotal

Summary

- **See computer output**