ANOVA - Factorial treatment arrangements and Randomized Block Design

Starting with the CRD, the linear model is

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

Source	*d.f.*	*EMS*
Treatment (between groups)	k-1	$\sigma^2 + n\sigma^2_\tau$
Error (within groups)	k(n-1)	σ^2
Total	kn-1	

The deviations within groups provides an estimate of the ε_{ij} part of the model, and this is our estimate of random error, σ^2. The τ_i term is estimate by the between groups variation. This is an estimate of the variance of the means between groups ($\sigma^2_{\bar{y}} = \sigma^2/n$) and, when multiplied by n, provides a second estimate of σ^2. If the null hypothesis (H_0:$\mu_1=\mu_2=\mu_3=\ldots=\mu_k$) is true then the between group estimate is only σ^2. However, if the null hypothesis is not true, then the between group variance contains a second term (σ_t^2), the variation due to the differences between the means (again, under the alternate hypothesis, H_1: some μ_i is different). This additional term is $n\sigma_t^2$, becomes $n\sigma_t^2$ after the previously mentioned multiplication by n. These calculations result in the two expected mean squares in the table above.

This model (CRD) makes no statement about the nature of the treatment. The treatment could be a simple selection of certain levels of interest (an *a priori* treatment arrangement), and these could be selected as either fixed or random effects. If random, the expected mean square for treatments is $\sigma^2 + n\sigma^2_\tau$, and if fixed they would be denoted as $\sigma^2 + n\frac{\sum_{i=1}^{n_i} \tau_i^2}{k-1}$ which is often noted as simply σ^2+Q_τ, where Q_τ is the fixed effect portion of the EMS. In either case, the F test is done as $F=^{MSTreatments}/_{MSError}$. If the null hypothesis is true, then the value of $n\sigma^2_\tau$ or Q_τ is not different from zero, then the F test should be near 1 and should not be "significant".

Factorial treatment arrangements.

	Treatment A	
Treatment B	**a1**	**a2**
b1	a1b1	a2b1
b2	a1b2	a2b2
b3	a1b3	a2b3

The treatments in a CRD could also be arranged in a "factorial" arrangement. In this case there would actually be two treatments which would be "cross classified" so that each level of treatment A would occur in combination with each level of treatment B. For example, suppose we have a 2 by 3 factorial with treatment A levels of a1 and a2, and treatment B levels of b1, b2 and b3. There are two ways in which this could be analized. It could be done as a simple CRD with 6 treatment "combinations" (a1b1, a1b2, a1b3, a2b1, a2b2, and a2b3) or it could be done as a "two-way ANOVA. If done as a CRD with 6 levels, the source table would be the same as for the example above.

If done as a two-way ANOVA, the model is still a CRD. The linear model and source table are as follows.

$$Y_{ijk} = \mu + \tau 1_i + \tau 2_j + \tau 1\tau 2_{ij} + \varepsilon_{ijk}$$

Source	*d.f.*	*EMS*
Treatment 1	t_A-1	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_2 n\sigma^2_{\tau 1}$
Treatment 2	t_B-1	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_1 n\sigma^2_{\tau 2}$
Treatment 1 by treatment 2 interaction	(t_A-1)(t_B-1)	$\sigma^2 + n\sigma^2_{\tau 1\tau 2}$
Error (within groups)	t_At_B(n-1)	σ^2
Total	t_At_Bn-1	

Note that the EMS for the treatments contain a component of variation for the treatment, and also contains a component for the interaction. Higher level treatments contain all the **random** terms from lower levels. Therefore, both treatments contain

the sampling error and the interaction term.

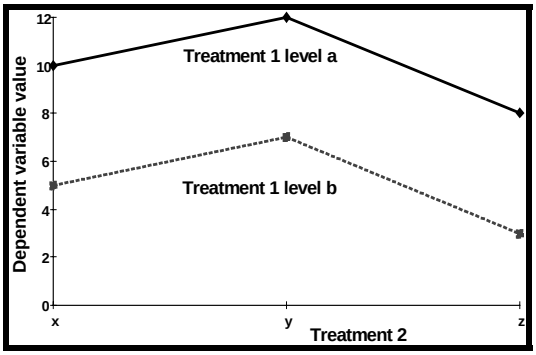
There is also a new source for the interaction itself. The interaction is a measure of the consistency or "additivity" of the two treatments in combination. For example, suppose the treatment a1 has a mean 5 units higher than the treatment a2 mean. This could mean that cell a1b1 is 5 higher than a2b1, that cell a1b2 is 5 higher than cell a2b2 and that cell a1b3 is 5 higher than cell a2b3.

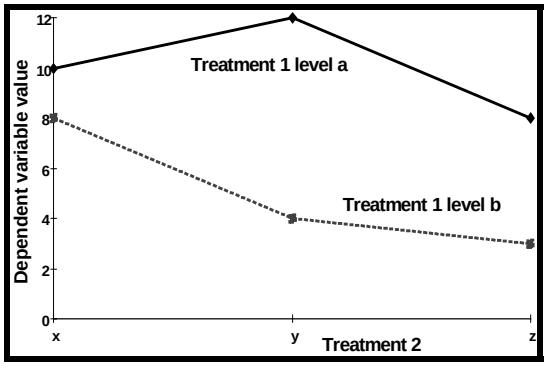
Case 1: no interaction	**a1**	**a2**
b1	10	5
b2	12	7
b3	8	3
marginal means	10	5

But what if we have the same marginal means, but the individual cells are not consistently different by 5 units. So the differences between the treatment levels "a" and "b" are the same (i.e. a is 5 units higher than b). However, the individual cells are not consistently higher, some are a little higher and others are a lot higher. This is an interaction.

Case 1: interaction	**a1**	**a2**
b1	10	8
b2	12	4
b3	8	3
marginal means	10	5

The easiest way to examine an interaction is to plot the means. The two sets of values above are plotted below. If there is no interaction the plots of the lines will appear to be parallel. If there is an interaction, the lines will not be parallel.





One last thing on factorials. Examine the EMS. We expect the F ratio to be 1 if the null hypothesis is true. If we wish to test H_0: $\sigma^2_{\tau 1}=0$, then we divide the EMS for treatment 1 by the appropriate error term.

However, if we calculate $F = \dfrac{\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_2 n\sigma^2_{\tau 1}}{\sigma^2}$, the test will differ from 1 if either the treatment component ($\sigma^2_{\tau 1}$) or the interaction compontent ($\sigma^2_{\tau 1\tau 2}$) is not equal to zero. Therefore, in order to do this test we must use the INTERACTION term as an error term. The correct F test is then given by

$$F = \frac{\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_2 n\sigma^2_{\tau 1}}{\sigma^2 + n\sigma^2_{\tau 1\tau 2}}.$$

This is true for RANDOM EFFECTS, However, if BOTH effects are fixed then the interaction is also

Source	***d.f.***	***EMS Random***	***EMS Fixed***
Treatment 1	t_1-1	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_2 n\sigma^2_{\tau 1}$	$\sigma^2 + t_2 n \frac{\sum_{i=1}^{t_1} \tau 1_i^2}{t_1 - 1}$
Treatment 2	t_2-1	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + t_1 n\sigma^2_{\tau 2}$	$\sigma^2 + t_1 n \frac{\sum_{i=1}^{t_2} \tau 2_i^2}{t_2 - 1}$
Treatment 1 by treatment 2 interaction	(t_1-1)* (t_2-1)	$\sigma^2 + n\sigma^2_{\tau 1\tau 2}$	$\sigma^2 + n \frac{\sum_{i=1}^{t_1 t_2} \tau 1\tau 2_i^2}{(t_1 - 1)(t_2 - 1)}$
Error (within)	t_1 t_2(n-1)	σ^2	σ^2
Total	kn-1		

fixed, and fixed effects do not occur in any source other than their own.

Randomized Block Design

The idea, setup and calculations behind RBD is similar to the factorial design with one large conceptual difference. Again we start with the CRD, and with the model $Y_{ij}=\mu+\tau_i+\varepsilon_{ij}$. In the randomized block design there is a new source of variation, but this new source does not come from subdividing the treatments as it does in the factorial design. The new source of variation comes from the error term.

For example, if we do an agricultural experiment where we apply 5 treatments to a field with 4 replicates of each treatment, then we need a field divided into 20 plots. However, if we can only fit 5 plots in a field (one per treatment) we could obtain the same number of replicates if we put the experiment into 4 different fields (4 fields each with one plot for for each of the 5 treatments). We still have 20 plots for 5 treatments. However, there is a new, potential source of variation, the variation between fields. If we ignore this variation it goes into the error term and we have what appears to be a CRD. However, the error is inflated because of between-field differences. We should instead factor out the "between field" differences as a new source of variation. These are like replicates and are called "blocks".

The linear model for the RBD is $Y_{ij}=\mu+\beta_i+\tau_j+\varepsilon_{ij}$. Here the error term, ε_{ij}, is also like an "interaction" term ($\tau\beta_{ij}$), but for block designs we consider the blocks to be like replicates, so the error term is a measure of random variation, σ^2.

If we have more than one plot per treatment in each field, then we have a second error term that is "within field" variation and the model is $Y_{ijk}=\mu+\beta_i+\tau_j++\beta\tau_{ij}+\varepsilon_{ijk}$. Both $\tau\beta_{ij}$ and ε_{ij} are error terms, one between blocks and the other within blocks. The expected mean squares are given below. The block effect interactions would always be considered random effects, the treatments may be either fixed or random. Note that the interaction between a fixed effect and a random effect is itself a random effect.

Source	***d.f.***	***EMS for Random Treatment, no within block replicates***	***EMS for Random Treatment, with within block replicates***	***EMS for Fixed Treatment, with within block replicates***
Block	b-1	$\sigma^2+tn\sigma^2_\beta$	$\sigma^2+n\sigma^2_{\beta\tau}+tn\sigma^2_\beta$	$\sigma^2+n\sigma^2_{\beta\tau}+tn\sigma^2_\beta$
Treatment	t-1	$\sigma^2+bn\sigma^2_\tau$	$\sigma^2+n\sigma^2_{\beta\tau}+bn\sigma^2_\tau$	$\sigma^2+n\sigma^2_{\beta\tau}+bn\frac{\sum_{i=1}^{t}\tau_i^2}{t-1}$
Treatment by block interaction	(t-1)(b-1)	σ^2	$\sigma^2+n\sigma^2_{\beta\tau}$	$\sigma^2+n\sigma^2_{\beta\tau}$
Error (within blocks)	tb(n-1)		σ^2	σ^2
Total	tbn-1			

Other examples of "blocks" would be "years", where an experiment is done several times in several years, or incubators if there was not enough space in a single incubator for the experimental material. Care must be taken that every treatment occurs in each block, preferably in equal numbers.

Nested designs

Sometimes some benefit can be obtained by sampling a site several times once it has been reached, or sampling a plot several times after a crop has been grown. For example, suppose we wish to compare the choloroflorocarbon content of air samples taken at high altitudes at various sites. The time and expense of going to a site, launching and recovering a baloon is great, so we may decide to launch our baloon several times at each site to give us a "within site" variability. Then there is the possibility of a failure of our sampling device (which will return with an air sample), so we may put several sample tanks on the baloon. Finally we get back to the laboratory with our samples, and we may make several determinations of choloroflorocarbon levels from each sampling tank.

When we are done, we have Sites which is what we want to compare (these are treatments). We have say 2 "baloon launches" within each site. Then we have 3 sampling tanks on each baloon (lets suppose than none fail). And finally we have, say, 4 measurements of cholоflorocarbon from each tank. The best measure of variation for "sites" is the within site launches, and this is also probably the largest variation because we are sampling different "air masses" with each launch. Then the second largest measure of variation is probably from the between tank (within launch) measure of variation. The smallest variation is probably measured by the "within tank" measurements.

Suppose that our objective is to measure the variability between sites, and that sites are randomly chosen. The analysis would be the same if sties were fixed. The expected mean squares are as follows.

Source	***d.f.***	***EMS for Random Treatment***
Sites	t-1	$\sigma^2 + n\sigma^2_\gamma + nb\sigma^2_\delta + nbl\sigma^2_\tau$
Launches(Site)	t(l-1)	$\sigma^2 + n\sigma^2_\gamma + nb\sigma^2_\delta$
Tank(Launch Site)	tl(b-1)	$\sigma^2 + n\sigma^2_\gamma$
Measure(Launch Site Tank)	tlb(n-1)	σ^2
Total	tlbn-1	

What is the appropriate error term for sites?

Have we gained anything by taking all of those extra samples?