

Summary

If the assumptions

- ▶ independence (holds by design in a CRD)
- ▶ Normal populations
- ▶ equal variances

hold, use

ANOVA F-test

If the assumptions **do not hold**

- ▶ use **Randomization/Permutation test**
- ▶ use **Non-parametric test** (see Section 3)

1.3 Multiple Comparison of Means

If the ANOVA F-test null hypothesis

$$H_0 : \mu_1 = \cdots = \mu_k$$

is **rejected**, then it is of interest to discover which of the means are different. For k groups, there are $c = k(k - 1)/2$ pairs of group means that can be compared.

Consider a “family” of hypothesis tests - a collection of tests of different hypotheses carried out independently on different data sets. For each test in the family, we consider testing the hypothesis at significance level α .

Notation

Label the tests $i = 1, \dots, c$, and for each i , label

- ▶ the null hypotheses H_{0i}
- ▶ the test statistics T_i
- ▶ the rejection regions \mathcal{R}_i

that are potentially different for each i .

We specify for each i ,

$$\alpha = P[T_i \in \mathcal{R}_i | H_{0i} \text{ is } \mathbf{TRUE}]$$

which implicitly defines \mathcal{R}_i . Note that α is the

“Test Type-I Error Rate” or “Comparisonwise Error Rate”

Now consider the results of all tests in the family; what is the “Familywise” Type-I error rate ?

Using the laws of probability

$$P[T_i \in \mathcal{R}_i | H_{0i} \text{ is } \mathbf{TRUE}] = \alpha$$

means that

$$P[T_i \notin \mathcal{R}_i | H_{0i} \text{ is } \mathbf{TRUE}] = 1 - \alpha$$

giving the probability that the test **does not reject** H_{0i} , if H_{0i} is in fact true, is $1 - \alpha$.

Now we consider all tests together;

$$P[\text{Each } T_i \notin \mathcal{R}_i | \text{Each } H_{0i} \text{ is } \mathbf{TRUE}] = (1 - \alpha)^c$$

This is the probability that each test results in the null hypothesis **not** being rejected, that is, the probability that we **never** commit a Type-I error.

Therefore the probability of **at least one** Type-I error is

$$\alpha_F = 1 - (1 - \alpha)^c$$

α_F is the **Familywise Error Rate**.

	$\alpha = 0.05$	$\alpha = 0.01$
c	α_F	α_F
5	0.226	0.049
10	0.401	0.096
50	0.923	0.395
100	0.994	0.634

Therefore, whenever we carry out a “family” of tests, we should not use the traditional choices of $\alpha = 0.05$ or 0.01 on each test.

The Bonferroni Method

To fix $\alpha_F = 0.05$, say, we need to use α on each test where

$$\alpha_F = 1 - (1 - \alpha)^c \iff \alpha = 1 - (1 - \alpha_F)^{1/c}$$

For example, if $\alpha_F = 0.05$ and $c = 10$, use

$$\alpha = 1 - (1 - 0.05)^{1/10} = 0.0051$$

It can be shown that

$$1 - (1 - \alpha)^c \approx c\alpha$$

Therefore, if α_F is the required **familywise error rate**, we must set the **comparisonwise error** rate to be $\alpha = \alpha_F/c$.

α_F/c is known as the **Bonferroni Correction**.

Confidence Intervals

For the $k = 2$ group comparison of means, a $100(1 - \alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}(n_1 + n_2 - 2)s_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where $t_{\alpha}(\nu_1)$ is the $1 - \alpha$ probability point of the Student-t distribution with ν_1 degrees of freedom (under the assumptions of independence, Normality and equal group variances).

If we move to a family of c tests, to get simultaneous confidence intervals for the differences in means $\mu_i - \mu_j$ for all pairs of i and j , we should adjust α to α_F when computing the $100(1 - \alpha)\%$ confidence interval.

SPSS gives twelve different methods for correcting the confidence interval for use in different experimental situations. For example

- ▶ *planned* comparisons $\mu_1 = \mu_3, \mu_7 = \mu_{10}$ etc.
- ▶ *all* comparisons

Three methods are recommended:

- ▶ **Tukey's** Method
- ▶ **Bonferroni's** Method
- ▶ **Scheffé's** Method

Having selected a multiple comparison correction method, we compute simultaneous confidence intervals for each comparison of means, and identify

- ▶ which means are significantly different
- ▶ the ranking of differences $\mu_i - \mu_j$ in terms of magnitude.

1.4 Randomized Block Designs

A **randomized block design** used **matched** experimental units organized into sets known as **blocks** and assigns one member from the set to each treatment.

For k treatments

1. Compile b blocks of k experimental units, with each block comprising units that are similar.
2. Assign one unit from each block to each treatment at random.

Then there are a total of $n = bk$ measured responses.

We wish to **compare treatments** whilst acknowledging that there may be **differences between the blocks**.

That is, the observed variation is due to

TREATMENTS and BLOCKS and ERROR

rather than merely

TREATMENTS and ERROR

as in the CRD.

Example (SAT Scores)

- ▶ **Response** : Measured SAT Score
- ▶ **Factor** : Sex
- ▶ **Factor-levels** : $k = 2$ (Female/Male)
- ▶ **Blocks** : $b = 5$ (Previous GPA, within same school)

i.e. $k = 2, b = 5 \therefore n = 10$.

	Block	Female SAT	Male SAT
1	A: 2.75	540	530
2	B: 3.00	570	550
3	C: 3.25	590	580
4	D: 3.50	640	620
5	E: 3.75	690	690

Example (SAT Scores (continued))

This design recognizes that GPA score and school are likely to explain some variation in SAT Score, **but that neither is directly related to the “treatment” of interest (SEX - Female/Male).**

i.e. the **blocking** variable removes systematic variation in response that is not of primary interest.

We pick one Female and one Male in each school/GPA category, and pair them.

Example (Treatment for Hypertension)

- ▶ **Response** : Blood Pressure (mgHg)
- ▶ **Factor** : Drug Type
- ▶ **Factor-levels** : $k = 3$ (Drug 1, Drug 2, Drug 3)
- ▶ **Blocks** : $b = 4$ Age/Sex combinations
 - ▶ Female/Under 50
 - ▶ Male/Under 50
 - ▶ Female/Over 50
 - ▶ Male/Over 50

i.e. $k = 3, b = 4 \therefore n = 12$.

Testing for Equal Treatment Means

An ANOVA F-test can be constructed for a RBD. Let

- ▶ $i = 1, \dots, k$ index **treatments**
- ▶ $j = 1, \dots, b$ index **blocks**

i.e. x_{ij} is the response for the i th treatment in the j th block. Let

- ▶ \bar{x}_i be the i th **treatment** mean
- ▶ $\overline{x_j^{(B)}}$ be the j th **block** mean
- ▶ \bar{x} be the **overall** mean

Let

$$SST = \sum_{i=1}^k b(\bar{x}_i - \bar{x})^2$$

$$SSB = \sum_{j=1}^b k(\overline{x_j^{(B)}} - \bar{x})^2$$

$$SS = \sum_{i=1}^k \sum_{j=1}^b (x_{ij} - \bar{x})^2$$

SST: Sum of Squares for **Treatments**

SSB: Sum of Squares for **Blocks**

SS: **Total** Sum of Squares

Finally

$$SS = SST + SSB + SSE \quad \therefore \quad SSE = SS - SST - SSB$$

SSE: Sum of Squares for **Errors**

Test statistic is

$$F = \frac{MST}{MSE}$$

where

$$MST = \frac{SST}{k - 1} \quad MSE = \frac{SSE}{n - b - k + 1}$$

ANOVA F-test to compare treatment means in a randomized block design

Theorem (ANOVA F-test for a RBD)

To test

$$H_0 : \mu_1 = \cdots = \mu_k$$

H_a : *At least one pair of treatment means different.*

use the test statistic

$$F = \frac{MST}{MSE}$$

*If H_0 is **TRUE***

$$F \sim \text{Fisher-}F(k - 1, n - b - k + 1)$$

- this defines the rejection region for significance level α , and the p-value, in the usual way.

Assumptions:

1. Experimental units (between blocks) are independent, and treatments are allocated at random (within blocks).
2. Normality
3. bk block/treatment combinations correspond to populations with equal variances.

ANOVA Table

SOURCE	DF	SS	MS	F
TMTS	$k - 1$	SST	MST	$F = MST/MSE$
BLOCKS	$b - 1$	SSB	MSB	
ERROR	$n - k - b + 1$	SSE	MSE	
TOTAL	$n - 1$	SS		

After the ANOVA test is complete, and the hypothesis

$$H_0 : \mu_1 = \cdots = \mu_k$$

is **rejected**, we can proceed with the “post-hoc” tests of hypotheses $\mu_i = \mu_j$ for $i \neq j$.

Notes:

1. In a RBD, it is not (in general) possible to estimate individual treatment means, that is, \bar{x}_i does not estimate μ_i as it is an average across blocks, which are believed to be different in terms of response.

2. **Testing the Block Means** However, we can test whether the block means $\mu_1^{(B)}, \dots, \mu_b^{(B)}$ are significantly different. For

$$H_0 : \mu_1^{(B)} = \dots = \mu_b^{(B)}$$

we use the F statistic

$$F = \frac{MSB}{MSE}$$

where

$$MSB = \frac{SSB}{b - 1}$$

If H_0 is **TRUE**

$$F \sim \text{Fisher-F}(b - 1, n - k - b + 1)$$

That is, we treat the blocks as levels of another factor, and test to see whether this factor affects response.

Example (Soil Analysis (see handout))

Results of two ANOVA F-tests:

Test of	F	p	Conclusion
SOLVENT	0.673	0.585	No Difference
SOIL	10.568	0.001	Difference

Here SOLVENT is the **treatment** variable, SOIL is the **blocking** variable.

- Remember to check the assumptions (independence, normality, equal variances in each treatment/block combination)

Equal variances may be hard to check as we only have one observation per treatment/block comparison.

Comment: The “sum of squares” decompositions

$$CRD : SS = [SST] + SSE$$

$$RBD : SS = [SST + SSB] + SSE$$

are both of the form

$$\begin{array}{l} \text{TOTAL} \\ \text{VARIATION} \end{array} = \begin{array}{l} \text{SYSTEMATIC} \\ \text{VARIATION} \end{array} + \begin{array}{l} \text{RANDOM} \\ \text{VARIATION} \end{array}$$

$$\begin{array}{l} \text{“SYSTEMATIC”} \\ \\ \text{“RANDOM”} \end{array} \quad \left\{ \begin{array}{l} \text{For the CRD: } SST \\ \text{For the RBD: } SST + SSB \end{array} \right.$$

$$\begin{array}{l} \text{“RANDOM”} \\ \\ \text{For both: } \end{array} \quad SSE$$

We have studied the

Randomized **Complete** Block Design

where each block/treatment combination has one experimental unit.

An incomplete design could also be considered, where some block/treatment combinations are omitted. However, this design does not lead to straightforward ANOVA analysis.

1.5 Factorial Experiments

Designs studied so far:

- ▶ CRD - one factor
- ▶ RBD - one factor, plus one blocking variable, so two factors in total, where one (the blocking variable) is a known source of systematic variation.

However, in the RBD, we must assume that the treatments behave in a similar way across blocks.

Let i index treatments ($1 \leq i \leq k$) and consider block j , and two treatment (factor levels) i_1 and i_2 .

In an RBD, we assume that

$$E[X_{i_1j} - X_{i_2j}] = \mu_{i_1} - \mu_{i_2}$$

which does **NOT** depend on j .

That is, the expected difference in response due to the two treatments does not depend on the block.

But perhaps the difference **does** depend on block; perhaps we have **INTERACTION**.

In the current RBD, we do not have enough data to look for this. We now seek to extend the RBD to allow for tests for interaction; we do this by using **replication**.

RBD with Balanced Replication

Suppose we have r observations per block/treatment combination (termed *replicates*), so that we have $n = bkr$ experimental units in total.

Balanced designs have equal numbers of replicates in each block/treatment combination.

In this design, all the quantities

$$SST, SSB, SSE, SS$$
$$MST, MSB, MSE$$

can be defined, and an ANOVA F-test can be carried out - the only difference is that $n = bkr$.

- ▶ Sum of Squares for Treatments (SST)

$$SST = \sum_{i=1}^k br(\bar{x}_i - \bar{x})^2$$

- ▶ Sum of Squares for Blocks (SSB)

$$SSB = \sum_{j=1}^b kr(\overline{x_j^{(B)}} - \bar{x})^2$$

- ▶ Overall Sum of Squares (SS)

$$SS = \sum_{i=1}^k \sum_{j=1}^b \sum_{t=1}^r (x_{ijt} - \bar{x})^2$$

and $SSE = SS - SST - SSB$

Third index t indexes the replicates.

The RBD with replication does allow the investigation of interaction. The new test is based on the decomposition

$$SS = SST + SSB + SSI + SSE$$

where SSI is the sum of squares for Interaction.

We have SST , SSB and SS as before, and

$$SSI = \sum_{i=1}^k \sum_{j=1}^b r(\bar{x}_{ij} - \bar{x}_i - \overline{x_j^{(B)}} + \bar{x})^2$$

where

$$\bar{x}_{ij} = \frac{1}{r} \sum_{t=1}^r x_{ijt} \quad i = 1, \dots, k, j = 1, \dots, b$$

is the sample mean for replicates in (i, j) th treatment/block combination.

Testing in the RBD with Replication

The three F statistics

$$F = \frac{MST}{MSE} \quad F = \frac{MSB}{MSE} \quad F = \frac{MSI}{MSE}$$

can be used to test for significant Treatment, Block and Interaction effects respectively.

Now

$$MSE = \frac{SSE}{\text{Error d.f.}}$$

But what is “Error d.f.” ? It is a constant that dictates how large SSE should be on average.

The general rule for computing the error d.f. for any model is

$$\text{Error d.f.} = n - p$$

where n is the total **sample size** and p is the total **number of parameters fitted**.

How many parameters do we fit ?

- ▶ No Interaction

$$p = 1 + (b - 1) + (k - 1)$$

that is, the overall mean μ , plus the $b - 1$ differences from μ due to the blocks, plus the $k - 1$ differences from μ due to the treatments.

- ▶ Interaction

$$p = bk$$

that is, one parameter in each cell of the two-way table of blocks by treatments.

Thus

- ▶ No Interaction

$$p = 1 + (b - 1) + (k - 1) = b + k - 1$$

parameters, so

$$\text{Error d.f.} = n - p = n - b - k + 1$$

- ▶ Interaction: we fit $p = bk$ parameters, so

$$\text{Error d.f.} = n - p = n - bk$$

It transpires that if

$$MSI = \frac{SSI}{(b-1)(k-1)}$$

is the Mean Square for Interaction, then

$$F = \frac{MSI}{MSE}$$

yields a test statistic suitable for testing interaction. If there is **no interaction**, then

$$F \sim \text{Fisher-F}((b-1)(k-1), n - bk)$$

where $n = bkr$.

Why $(b-1)(k-1)$? This is the number of **extra** parameters we fit to include the interaction.

For the CRD:

$$\begin{array}{ccc} & H_a & H_0 \\ \text{FULL MODEL} & \longrightarrow & \text{NULL MODEL} \\ k \text{ parameters} & \longrightarrow & 1 \text{ parameter} \end{array}$$

so there are $(k - 1)$ extra parameters, and SST varies on $(k - 1)$ degrees of freedom.

For the RBD: the (i, j) th treatment/block combination has mean

$$\mu_i + \mu_j^B$$

so for testing for a TREATMENT effect

$$\begin{array}{ccc} H_a & & H_0 \\ \text{FULL MODEL} & \longrightarrow & \text{NULL MODEL} \\ k \text{ parameters} & \longrightarrow & 1 \text{ parameter} \end{array}$$

so there are $(k - 1)$ extra parameters, and SST varies on $(k - 1)$ degrees of freedom.

$$\mu_1, \dots, \mu_k \longrightarrow \mu$$

For testing for a BLOCK effect

$$\begin{array}{ccc} & H_a & H_0 \\ \text{FULL MODEL} & \longrightarrow & \text{NULL MODEL} \\ b \text{ parameters} & \longrightarrow & 1 \text{ parameter} \end{array}$$

so there are $(b - 1)$ extra parameters, and *SSB* varies on $(b - 1)$ degrees of freedom.

$$\mu_1^{(B)}, \dots, \mu_k^{(B)} \longrightarrow \mu^{(B)}$$

These models and tests can be fitted and carried out even if we do not have replication.

With replication, we can investigate the interaction, that is the model where the (i, j) th treatment/block combination has mean

$$\mu_i + \mu_j^B + \mu_{ij}$$

rather than the model where

$$\mu_i + \mu_j^B$$

that is, we wish to test

$$H_0 : \mu_{ij} = 0 \quad \text{for all } i \text{ and } j$$

$$H_a : \mu_{ij} \neq 0$$

In the **full interaction** model: we fit bk parameters

In the **restricted, no interaction** model: we fit

$$1 + (b - 1) + (k - 1) = b + k - 1$$

parameters. Therefore the difference is

$$bk - (b + k - 1) = bk - b - k + 1 = (b - 1)(k - 1)$$

and SSI varies on $(b - 1)(k - 1)$ degrees of freedom.

ANOVA Table

SOURCE	DF	SS	MS	F
TMTS	$k - 1$	SST	MST	F_T
BLOCKS	$b - 1$	SSB	MSB	F_B
INTERACTION	$(b - 1)(k - 1)$	SSI	MSI	F_I
ERROR	$(n - bk)$	SSE	MSE	
TOTAL	$n - 1$	SS		

where

$$MST = \frac{SST}{k - 1} \quad MSB = \frac{SSB}{b - 1}$$

$$MSI = \frac{SSI}{(b - 1)(k - 1)} \quad MSE = \frac{SSE}{n - bk}$$

and

$$F_T = \frac{MST}{MSE} \quad F_B = \frac{MSB}{MSE} \quad F_I = \frac{MSI}{MSE}$$

Example: Batteries Data (see handout)

Dependent Variable: Battery Life					
Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	59,154.000	8	7,394.250	11.103	0.000
Intercept	398,792.250	1	398,792.250	598.829	0.000
material	10,633.167	2	5,316.583	7.983	0.002
temp	39,083.167	2	19,541.583	29.344	0.000
material * temp	9,437.667	4	2,359.417	3.543	0.019
Error	17,980.750	27	665.954		
Total	475,927.000	36			
Corrected Total	77,134.750	35			

R Squared = .767 (Adjusted R Squared = .698)

For $\alpha = 0.05$, there is a significant **temp** effect ($p < 0.001$), and a significant **material** effect ($p = 0.002$), and a significant interaction ($p = 0.019$)

NB: If we do not have replication, we CANNOT fit the interaction.
Recall that

$$\text{Error d.f.} = n - bk$$

but if $r = 1$, $n = rbk = bk$, so the error d.f. is zero.

In fact, $SSE = 0$ also, so the MSE is not defined.

We now study multifactor designs, to assess the effects and interactions of several factors simultaneously.

We consider all possible combinations of

FACTOR A with a levels

FACTOR B with b levels

FACTOR C with c levels

⋮

to define the treatments in a factorial design.

Factorial Experiments

A **complete factorial experiment** is one in which every combination of a number of factors is utilized.

i.e. the number of treatments is equal to the total number of factor-level combinations.

We focus on two factor experiments

FACTOR A with a levels

FACTOR B with b levels

so there are ab treatments in total.

A two-way layout with $a = 3$ and $b = 5$.

		Factor B				
		1	2	3	4	5
Factor A	1					
	2					
	3					

This design is very similar to the RBD, but now the second factor is not a blocking factor;

- ▶ that is, the ab treatment populations are constructed independently from the same base population, or from populations not necessarily believed to be systematically different.
- ▶ individuals from the same base population are assigned at random to one of the ab treatments.

In this design we can study the effect of Factor A and Factor B (**main effects**) as well as the **interaction** provided we have (balanced) replication.

We construct ANOVA F-tests based on the decomposition

$$SS = SST_A + SST_B + SSI_{AB} + SSE$$

- ▶ Sum of Squares for Treatments due to factor A (SST_A)

$$SST_A = \sum_{i=1}^a br(\bar{x}_{i.} - \bar{x}_{..})^2$$

- ▶ Sum of Squares for Treatments due to factor B (SST_B)

$$SST_B = \sum_{j=1}^b ar(\bar{x}_{.j} - \bar{x}_{..})^2$$

- ▶ Sum of Squares for Interaction (SSI_{AB})

$$SSI_{AB} = \sum_{i=1}^a \sum_{j=1}^b r(\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$$

New notation:

- ▶ sample mean for Factor A level i

$$\bar{x}_{.i} = \frac{1}{br} \sum_{j=1}^b \sum_{t=1}^r x_{ijt} \quad i = 1, \dots, a$$

- ▶ sample mean for Factor B level j

$$\bar{x}_{.j} = \frac{1}{ar} \sum_{i=1}^a \sum_{t=1}^r x_{ijt} \quad j = 1, \dots, b$$

- ▶ sample mean for replicates in (i, j) th factor combination

$$\bar{x}_{ij} = \frac{1}{r} \sum_{t=1}^r x_{ijt} \quad i = 1, \dots, a, j = 1, \dots, b$$

- ▶ overall sample mean

$$\bar{x}_{..} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b \sum_{t=1}^r x_{ijt}$$

These allow computation of

$$SST_A, SST_B, SSI_{AB}, SS, SSE$$

$$MST_A, MST_B, MSI_{AB}, MSE$$

using the degrees of freedom identical to those in the RBD with replication.

Tests for

- ▶ significant effect for Factor A
- ▶ significant effect for Factor B
- ▶ significant interaction

will be carried out as before using an ANOVA table.

ANOVA Table

SOURCE	DF	SS	MS	F
FACTOR A	$a - 1$	SST_A	MST_A	F_A
FACTOR B	$b - 1$	SST_B	MST_B	F_B
INTERACTION	$(a - 1)(b - 1)$	SSI_{AB}	MSI_{AB}	F_{AB}
ERROR	$(n - ab)$	SSE	MSE	
TOTAL	$n - 1$	SS		

If Factor A is not influential (H_0 specifying no difference between responses at different levels of factor A), then

$$F_A \sim \text{Fisher-F}(a - 1, n - ab)$$

Similarly,

No effect of Factor B : $F_B \sim \text{Fisher-F}(b - 1, n - ab)$

No Interaction : $F_{AB} \sim \text{Fisher-F}((a - 1)(b - 1), n - ab)$

SEE EXAMPLES HANDOUT

Note: For two factors A and B, the **main effects plus interaction** model can be written

$$A + B + A.B$$

whereas the **main effects only** can be written

$$A + B$$

The models

$$A + A.B$$

$$B + A.B$$

do not make sense.

For a two factor design, the only models that should be considered and or reported are

MODEL	FACTOR	INTERACTION
NULL	NONE	NONE
A	A	NONE
B	B	NONE
A+B	A,B	NONE
A+B+A.B	A,B	YES

that is, if the interaction is significant, the only model you should report is

$$A + B + A.B$$

Note: ANOVA analysis for the RBD and FD (both with replication) are identical. The only difference lies in the interpretation of the factors

- ▶ RBD: one blocking, one treatment factor
- ▶ FD: two treatment factors

“Blocking” factors are known or strongly believed to have a significant effect on the response.

Estimating Effect Size

In multifactor designs, parameter estimation can be carried out in different parameterizations

For the CRD (one-way layout):

- ▶ Natural parameters: μ_1, \dots, μ_k
- ▶ Contrast parameters: $\beta, \beta_0, \dots, \beta_{k-1}$ where

$$\beta = \mu_k \quad \beta_i = \mu_i - \mu_k, \quad i = 1, \dots, k - 1$$

that is, differences from baseline.

For the two-factor designs (RBD/FD): In the two-way layout, with cells (i, j) , $i = 1, \dots, a$, $j = 1, \dots, b$. The cell means are m_{ij} , where

$$m_{ij} = \mu_{i.} + \mu_{.j} + \mu_{ij}$$

where $\mu_{i.}$ gives the Factor A contribution, $\mu_{.j}$ gives the Factor B contribution, and μ_{ij} gives the interaction.

The parameterization used by SPSS is the contrast parameterization is

$$\begin{aligned}m_{ij} &= \beta_0 & i = a, j = b \\&= \beta_0 + \beta_i^{(A)} & i = 1, \dots, a - 1, j = b \\&= \beta_0 + \beta_j^{(B)} & i = a, j = 1, \dots, b - 1 \\&= \beta_0 + \beta_i^{(A)} + \beta_j^{(B)} + \gamma_{ij}^{(AB)} & i = 1, \dots, a - 1 \\& & j = 1, \dots, b - 1\end{aligned}$$

where

$$\begin{aligned}\beta_i^{(A)} &: \text{contrasts for factor A} \\ \beta_j^{(B)} &: \text{contrasts for factor B} \\ \gamma_{ij}^{(AB)} &: \text{interaction}\end{aligned}$$

SPSS takes the a th level of factor A and the b th level of factor B as the baseline, and looks at differences compared to this baseline.

The ab parameters are

β_0	1
$\beta_1^{(A)}, \dots, \beta_{a-1}^{(A)}$	$(a - 1)$
$\beta_1^{(B)}, \dots, \beta_{b-1}^{(B)}$	$(b - 1)$
$\gamma_{ij}^{(AB)}, i = 1, \dots, a - 1, j = 1, \dots, b - 1$	$(a - 1)(b - 1)$
Total	ab

For example: $a = 3, b = 4$.

		Factor B			
		1	2	3	4
Factor A	1	①	②	③	$\beta_0 + \beta_1^{(A)}$
	2	④	⑤	⑥	$\beta_0 + \beta_2^{(A)}$
	3	$\beta_0 + \beta_1^{(B)}$	$\beta_0 + \beta_2^{(B)}$	$\beta_0 + \beta_3^{(B)}$	β_0

where

$$\textcircled{1} = \beta_0 + \beta_1^{(A)} + \beta_1^{(B)} + \gamma_{11}^{(AB)}$$

$$\textcircled{6} = \beta_0 + \beta_2^{(A)} + \beta_3^{(B)} + \gamma_{23}^{(AB)}$$

and so on.

Estimation is still straightforward:

PARAMETER	ESTIMATE
β_0	\bar{x}_{ab}
$\beta_i^{(A)}$	$\bar{x}_{i.} - \bar{x}_{ab}$
$\beta_j^{(B)}$	$\bar{x}_{.j} - \bar{x}_{ab}$
$\gamma_{ij}^{(AB)}$	$\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{ab}$

for $i = 1, \dots, a, j = 1, \dots, b$.

Other parameterizations can be used.

Final Note on ANOVA

We have studied the simplest design scenarios: extension to

- ▶ *incomplete*
- ▶ *unbalanced*
- ▶ *nested*
- ▶ *random effect*

designs are possible.

Furthermore SPSS has greater functionality: for example, it has the capability to carry out ANOVA-like analyses even for the case of non-equal variances (when Levene's test rejects the hypothesis of equal variances).