Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

https://doi.org/10.23913/ride.v15i29.2134

Scientific articles

Esperanza del cuadrado medio. Una explicación didáctica

Hope of the middle square. A didactic explanation

Esperança do quadrado médio. Uma explicação didática

Cienfuegos Velasco María de los Ángeles
Universidad Autónoma del Estado de México, Unidad Académica Profesional Chimalhuacán, México
angelescien@hotmail.com
https://orcid.org/0000-0002-8423-8088

Resumen

La Esperanza del Cuadrado Medio (ECM) es un referente que se considera importante en el análisis de varianza; su utilidad radica en el análisis de variaciones entre grupos en diseños de investigación vía experimento. Dicho análisis debe considerar las nociones de factor fijo, aleatorio, cruzado o anidado, así como los modelos identificados. Esto permitirá generar la ECM que, a su vez, ayudará a realizar los contrastes de hipótesis necesarios para resolver el modelo. De esta manera, decidir sobre el rechazo o no de la hipótesis planteada es clave en las conclusiones que se presenten en la investigación que se desarrolla. Es un estudio documental que explora y explica el proceso de la ECM y responde a la pregunta ¿cómo ofrecer una explicación didáctica a través de exposición de reglas y sus aplicaciones para el cálculo de la ECM? Se concluye haciendo énfasis en la importancia de seguir recomendaciones y reglas para obtener la ECM.

Palabras Claves: Esperanza de los Cuadrados Medios, aleatorio y modelos aleatorios.

Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

Abstract

The Mean Square Expectancy (ECM) is areference that is considered important in the análisis of variance; its usefulness lies in the analysis of variations between groups in experimental research designs. Said analysis must consider the notions of fixed, rondom, crossed or nested factor, as well as the identified models. This will allow the generation of the ECM which, in turn, will help to carry out the hypothesis test necessary to salve the model. In this way, deciding whether or not to reject the proposed hypothesis is key to the conslusions presentad in the research that is developed. It is a documentary study that explores and explains the MSE process and answers the question: how to offer a didactic explanation through the presentation of rules and their applications for the calculation of the ECM? It concludes by emphasizing the importance of following recommendations and rules to obtain the ECM.

Keywords: Expectancy of Mean Squares, random and random models.

Resumo

A Expectativa Quadrática Média (EQM) é uma referência considerada importante na análise de variância; Sua utilidade reside na análise de variações entre grupos em projetos de pesquisa experimental. Esta análise deverá considerar as noções de fator fixo, aleatório, cruzado ou aninhado, bem como os modelos identificados. Isto permitirá a geração do EQM que, por sua vez, ajudará a realizar os testes de hipóteses necessários à resolução do modelo. Desta forma, decidir se rejeita ou não a hipótese proposta é fundamental para as conclusões apresentadas na investigação que se desenvolve. É um estudo documental que explora e explica o processo de EQM e responde à questão: como oferecer uma explicação didática através da apresentação de regras e suas aplicações para o cálculo da EQM? Conclui enfatizando a importância de seguir recomendações e regras para obter a EQM.

Palavras-chave: Expectativa de Quadrados Médios, modelos aleatórios e aleatórios.

Date of Reception: February 2024 **Acceptance Date:** August 2024

Introduction

In this paper, the topic of Expectation of the Mean Squares (ECM) is developed , which requires previous learning of the statistical model with all its implications: concept of population or universe, management of variables, measurement scales, assumptions of the model such as normality, homogeneity of variances, independence of errors; in itself, they are statistical knowledge related to the research methodology and, in this case, the ECM is used, in particular, in experimental research.

It is specified that the ECM has been used almost exclusively by geneticists; this is because in genetics the presence of random variables or factors is very frequent in their models. Its use is required in certain special circumstances; for example, when sampling is done in the experimental units of a randomized complete block design and also, when in certain circumstances that arise when describing the inter-block information recovery process in incomplete block designs (Martínez, 1988).

The above is stated because the sample observations in the experimental units are not randomized; that is, the samples in the experimental units are not independent, which has repercussions on the structure of the models, such as the generation of a Restriction Error (REE); that is, the model presents a restriction for its proper functioning of randomizing the blocks and, as they are not randomized, this REE arises and of course the ECM works to see the variation between the groups; ignoring it leads to not giving an adequate interpretation of results and not knowing which hypothesis to reject and which not to reject. (Martínez, 1988).

The use of is la ECMnot only due to the presence of random factors in the model, but mainly because the treatments, due to their systematic location within the blocks, for some reason have not been randomized and also because the blocks are not randomized either, which occurs mainly in the agricultural sector and also in other areas of knowledge such as animal husbandry, biology, medicine, and industry. However, the conception that the blocks are randomized is erroneously carried over from Fisher (1936).

Consequently, the presence of random factors and the non-randomization of one or more of these factors, including blocks, leads in both cases to the use of the EdeR la ECMand to pseudo-experimental research projects (observational and comparative). This allows for adequate hypothesis testing (knowing what to reject and what not to reject).

When treatments have been properly randomized, the EofR is not necessary. The manipulation or transformation of the research material is also necessary for the experiment

to exist. However, even if the latter occurs, without randomization the experiment ceases to be one.

Do geneticists and researchers in general consider this situation when they do research? Why can they be working with experiments when in reality they are not? In how many academic theses are these techniques used and in how many are they not used when they should be used? It is necessary to be aware of this situation.

Fundamental methodological concepts

It is necessary that methodology and statistics be taught in a dependent manner, depending on the type of research. Advisors, teachers, methodologists, statisticians and researchers in general must work on statistical techniques and their application to methodological procedures, thereby correctly applying methodological techniques to statistics. Therefore, the following is briefly pointed out:

1. Practice the proper use of statistics and research in general and, in particular, research via experiment (Table 1), and so on in the rest of the ten types or research projects (Table 2), with the methodological steps that are required.
2. Proper handling of randomization and blocks.
3. Proper management of EdeR, NDE, confounding factor, hypothesis testing and many other classical techniques that are very frequently presented in the field of scientific research.

Some methodological and statistical applications and clarifications inherent to theses and research projects are derived from the content of what is called the scientific research matrix (Tables 2 and 3), based on four dichotomous criteria (Tables 1 and 2), which when combined give rise to the ten types or projects of scientific research, topics that are briefly described below.

Table 1. Four criteria by which scientific research is classified

Criterion number	Dichotomous criterion	Characteristics that define the criterion
1	Observational-Experimental What defines the researcher	Absence – presence (of randomization and manipulation of the research material)
2	Prospective-Retrospective	Present and future – Past
3	Transversal-Longitudinal It is defined by the researcher	One measurement - multiple measurements (evolution, monitoring of the phenomenon)
4	Monogroup - Comparative	A population or group – More than one group

Source: (Mendez, 1984)

Table 2. Scientific research matrix.

Combination of the four criteria for classifying research into ten types of scientific research design studies or projects (common name).					
Dichotomous classification criteria					
1	2	3	4	According to table 1	
Observational either Experimental	Prospective either Retrospective	Longitudinal either Cross	Monogroup either Comparative	Design, study or project. (Common name).	Project and clue
Observational	Prospective	Cross	Monogroup	Survey. Single group	1 D -C
Observational	Retrospective	Cross	Monogroup	Survey. Single group	2 D -C
Observational	Prospective	Cross	Comparative	Comparative Survey	3 D -B
Observational	Retrospective	Cross	Comparative	Survey. Comparative	4 D -B
Observational	Retrospective	Longitudinal	Monogroup	Case review	5 C
Observational	Retrospective	Longitudinal	Comparative Effect-cause	Cases and controls	6-B
Observational	Retrospective	Longitudinal	Comparative Cause-effect	Historical perspective	7 B
Observational	Prospective	Longitudinal	Monogroup	A cohort	8 C
Observational	Prospective	Longitudinal	Comparative	Several cohorts	9- B
Experimental	Prospective	Longitudinal Cross	Comparative	Experiment	10-A

Source: Mendez (1984)

Protocol 1 and 2 are both D (survey) and C (no experiment, no pseudo experiment).

Protocols 3 and 4 are both D (survey) and B (pseudo experiment).

Protocol 5 and 8 is C (not experiment or pseudo experiment).

Protocol 6,7 and 9 are B (pseudo experiment)

Protocol 10 is A (experiment).

Table 3. Subdivision of the scientific research matrix.

According to table 2: dichotomous monogroup-comparative criterion and observational-experimental criterion; a new classification of research results. Four new projects:		
New classification of the Investigation. New type of projects.	Quantity of each classification. Type of criterion.	Number of project(s) and key(s)
1. Experiment (A).	Comparatives: 1	10-A
2. Pseudo experiment (B).	Comparatives: 5	3, 4, 6, 7, 9-B
3. No experiments or pseudo experiments (C)	Monogroup: 4	1, 2, 5, 8- C
4. Surveys (D).	Monogroup: 2 Comparative: 2	1, 2- D 3, 4-D
	Total: 14	Total: 14

Fountain: Cienfuegos (1990)

There are 14 projects in this new classification (Table 3), instead of 10 (Table 2), because:

- A project is an experiment (A), and comparative.
- Five Projects are pseudo experiments (B) and, in addition, comparative.
- There are four, not experiments or pseudo experiments (C), also monogroup
- Two are single-group surveys (D) and two are comparative.

Now, if research leads to research, then experimentation leads to experimenting. Let's see: The first (research) is present throughout Table 2 (the ten projects), whose common name is in column 5. The second (experimentation) is at the bottom of Table 2 (the experiment). That is to say, when an experiment is carried out, research is done; however, research may or may not be done with experiments. It is important to differentiate the concept of research from that of experimentation and to apply the latter to cases in which the project is truly an experiment.

The experiment must meet the requirement of randomizing treatments and manipulating, transforming or modifying the research material. Randomizing means giving each participant or treatment the same chance of being included in the experiment. It should be clear that the experimental refers to the criterion, the experiment to the name of the project and the observational refers to the criterion, resulting in nine types of projects.

If there is no randomization, there is no experiment. This leads to the observational criterion and in particular (as a comparative study) to the pseudo experiment; it also leads to the EdeR, NDE and, consequently, to the way of carrying out hypothesis tests, to know what to reject, what not to reject. Randomization is an inseparable characteristic of the experiment

.

On the other hand, only in comparative projects (two or more populations) are confounding factors present. Its counterpart is the monogroup criterion. Both criteria (comparative and monogroup) are defined and determined by the researcher.

Other authors call the monogroup criterion the descriptive criterion, interpreting it as the fact of studying a population, when its most important function is to describe. In this regard, the descriptive is not exclusive to a group, to a population, comparative projects are also described. On the other hand, the monogroup criterion (of a single group) explains itself. The above, in a brief and summarized form, constitutes the necessary and sufficient raw material *for* the reader to assimilate the ECM subject, whose rules are discussed below.

Rules for calculating the ECM: Expected Values (EV), Variance Components (CV)

Easy rules are presented to calculate ECM, which has so far been common in research practice, the use of fixed models, to test hypotheses by the mean squares ratio: Ratio of the mean square of the factor under hypothesis (as numerator), to the mean square of the error of the complete model (as denominator). Relationship called calculated F (F_c). Mathematically it is expressed as follows:

$$Fc = \frac{\text{CM (of the factor under hypothesis) CM/Ho Gl . (numerator)}}{\text{CME (of the full model) CME Gl . (denominator)}} = \text{---------} \sim F_{tables}$$

Fc value is compared with the F value of tables , entering the degrees of freedom of the numerator, degrees of freedom (Gl) of the *denominator* and a certain level *of* significance alpha (α), which does not necessarily have to be 5% or 1%, as is customary, because it could be, according to the needs of the phenomenon studied, 4%, 7%..., to then use the following decision rules:

Yes, $Fc \geq F_{tables} \Rightarrow$Reject Ho: $\tau_1 = \tau_2 = \tau_3 = ... = \tau_t = \tau$

Yes, $Fc < F_{tables} \Rightarrow$Reject Ho: $\tau_1 = \tau_2 = \tau_3 = ... = \tau_t = \tau$

These rules are only valid for standard or classical models (fixed effects), provided that the variation factor under Hypothesis has been randomized; *There* may be two cases:

a) A fixed, random or mixed effects model is built, regarding the presence or absence of random factors (researcher's decision).
b) The model is built (based on the needs of the phenomenon studied) with one or more non-randomized factors. Pseudo experiments then arise (observational and comparative), which require the injection of the EdeR into the model.

In both cases it is necessary to calculate ECM, in order to highlight and make visible the test of the hypotheses involved in the research and, in the second case, also to inject one or more restriction errors into the model.

The researcher must be aware, from the planning of his research, of the situations a and b mentioned, which help him define the research design, statistical-mathematical model, techniques and statistical tests.

Specific case of the completely randomized design (CaA):

This is a very special case, when the treatments have not been randomized (systematically placed). This situation frequently occurs in agro-ecological, biological, agricultural, livestock, forestry, medical and industrial studies, among others. Non-randomized treatments are often wrongly interpreted as true and independent repetitions (Martínez, 1988). They are not because they are correlated, originating pseudo experiments or quasi experiments, with EdeR; the calculation of the column (ECM) is required to carry out hypothesis tests. To carry out research, the following is recommended:

a) Define, based on theory (deep knowledge of the phenomenon studied), the representative model of the population.
b) Define the analysis of variance (AvA) table. Initially, with the columns of source of variation (SV), degrees of freedom (DF), mean square (MS) and ECM. How to define ECM for each source of variation?

   It is recommended to write, according to Santizio (1974), the following:

   - The name: AdeV Table.
   - Design Type: Common name, standard or not.
   - Indicate levels or treatments for each variation factor.
   - Accompany each term of the model with the following:

     Below the term: The symbol of its corresponding CV.

     Above the term: Indicate whether it is fixed or random, as well as the number that corresponds to it, according to the term number of the model (in a circle).

For educational purposes, it is recommended that inexperienced researchers practice these sections. However, it is preferable to use them in all cases, in order to avoid making mistakes in the theoretical handling of the research. In addition:

- Write the symbolic subscripts: i, j, k… and the real ones: a, b, r...
- Describe for each thesis or project the assumptions of the model:

a) Type of data distribution
b) Homogeneity of data
c) Independence from errors.

- Write the name of the dependent variable, usually symbolized as Y, on the left, at the top of the table.
- If possible (for educational purposes), present the graph of the model.

Model 1. Random bifactorial (α, $\beta$and their interaction $\alpha\beta$.), to calculate ECM. Note in table 1: All coefficients of variation (CV) remain. Why? Because the variation factors are all random.

Table 4. The AdeV of Model 1 Bifactorial randomized Completely randomized design (CaA).

Experimental research (via experiment), with two variation factors.

			Random.	Random.		
Model			②	③	④	①	
$Y_{ijk} =$	$\mu+$	$\alpha_i +$	$\beta_j +$	$(\alpha\beta)_{ij}$	$+ \varepsilon_{k(ij)}$	(1)
		σ_i^2	σ_j^2	σ_{ij}^2	σ_ε^2	
	$\varepsilon_{k(ij)} \sim NI(0, \sigma_\varepsilon^2)$					

Source: Own elaboration

i = 1, 2, 3...a, levels of αrandom factor $_i$
j = 1, 2, 3...b, levels of the βrandom factor $_j$
k = 1, 2, 3...r, repetitions of each combination of treatments (balanced case).

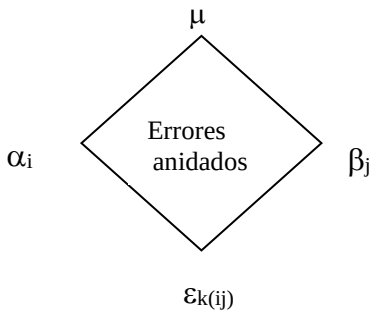
i, j, k are the symbolic subscripts
a, b, r, are the real subscripts

Table 5. For the Random Bifactor Model

Dependent variable Y			
FV	GL	CM	ECM
Factor α_i (Random)	(a-1)	CM_α	① ② ④ $\sigma_\varepsilon^2 + rb\ \sigma_i^2 + r\sigma_{ij}^2$
Factor β_j (Random)	(b-1)	CM_β	① ③ ④ $\sigma_\varepsilon^2 + ra\ \sigma_j^2 + r\sigma_{ij}^2$
Int Factor : $(\alpha\beta)_{ij}$	(a-1) (b-1)	$CM_{\alpha\beta}$	① ④ $\sigma_\varepsilon^2 + r\sigma_{ij}^2$
Error: $\varepsilon_{k(ij)}$	Ab (r-1)	CM_ε	σ_ε^2 ①
Total: Y_{ijk}	Apr-1		

Source: Own elaboration

Figure 1. Representation of the model, where α_i , β_j , are cross factors.



Source: Own elaboration

Description of the rules: When reading these rules, it is necessary to keep in mind the model (1). It will be assumed (in all examples) that everything is randomized (without EdeR). The application of the rules to calculate ECM is exemplified in principle with a random bifactorial model, which means that both factors and their interaction (when it exists) are random:

- Rule 1. In the AdeV, column ECM, the coefficient of variation (CV) is written σ_ε^2 as the first term for all sources of variation. In the model in question, the error $\varepsilon_{k(ij)}$, (epsilon k, within ij), is random and hierarchical or nested in ij .
- Rule 2. For each source of variation, the CVs of the model are written in the ECM column with the following condition: The subscript(s) of the factor in question, for example, subscript i of $\alpha_{i,}$ must be in said CVs.
- Rule 3. Automatically, according to rule 2, the subscribers that are not included are defined for each selected CV, whose values are placed as coefficients of said components, in addition to the coefficient r (number of repetitions), which are nothing more than the sample sizes (number of levels or treatments); that is, the sample sizes are replaced by their real value when carrying out the corresponding hypothesis tests.
- Rule 4. In the ECM column, not all the CVs thus selected are in the row corresponding to the variation factor in question; that is, the CVs may or may not be in that row. According to this rule (of the components that should or should not be in the corresponding row), to find out, the following criteria are applied. (See Table 5).
  - ✓ For example, in the row for factor α_i , and located (in the ECM column), in each of the corresponding CVs. Thus, in the CV, σ_{ij}^2 of the subscribers ij , I

ignore i (because it is in the row for factor α_i), leaving j; in such a way that if j is fixed the term disappears; but if j is random, the term remains. The same criterion applies to the rest of the CVs. It is a practical rule, which can be followed with interest.

- ✓ Alternative rule to rule 4, even more practical than the previous one: the CVs of the model whose subscript(s) exactly match that of the variation factor in question, remain in the corresponding row (whether fixed or random). This is the case of variation factor α_i and β_j .

Note that in Table 5, the CVs of each FV are left (because all sources of variation are random). This example encompasses or compresses all cases. To apply rules, focus on the following:

1. In the terms of the model.
2. In the symbolic subscripts (i, j, k...), for the model in question.
3. In the real subscripts (a, b, r...), because they take a certain value.
4. In the columns; FV, GL, F_{tables} and ECM.
5. In the CVs corresponding to each term of the model.
6. In the rows containing the CVs in the ECM column.
7. It is convenient to assign a number to the terms of the AdeV model and table, giving the number 1 to the error and the following to the rest of the terms from left to right.
8. Define for each term of the model, those that are fixed and those that are random.

Random bifactorial (α, $\beta$and interaction $\alpha\beta$), to apply rules and calculate ECM:

Application of rule 1: In all rows (already explained), the error is written as the first CV σ_ε^2

Application of rule 2: Which defines the CVs for the corresponding row: The application of this rule is exemplified (for each variation factor), based on the AdeV table (Table 5), as follows:

- In the first row: The variation factor α_i, whose subscriber is i, appears in the terms ②and ④of the model, whose CV, σ_i^2 and σ_{ij}^2 are written in the row of said factor, in addition to the CV, σ_ε^2

- In the second row: The variation factor β_j whose subscriber is j, appears in the terms ③and ④of the model, whose CVs are σ_j^2 and σ_{ij}^2, are written in the row of said factor, in addition to the CV σ_ε^2
- In the third row: In the interaction factor $(\alpha\beta)_{ij}$, the subscribers ij appear only in the term ④of the model, whose CV σ_{ij}^2 is written (for this reason) in said row, in addition to the CV σ_ε^2
- In the fourth row *:* The subscribers k(ij) of the error appear only in the ①model term, whose CV σ_ε^2 is written in the row corresponding to the error.

Rule 3 application . Which defines (in the ECM column), for each row, the coefficients of each CV. One of them is always "r", except for σ_ε^2

How to proceed for each variation factor:

- In the first row , corresponding to the variation factor α_i : In the CV, r σ_{ij}^2 all subscribers are present, so there are no coefficients to write (only r).
- In the same row, of the factor α_i , The CV rb σ_i^2 has as coefficient rb. Why? Because in said CV there is not the subscriber j = 1, 2, 3...b, this level b being the one that is placed as coefficient, in addition to r; that is: rb .
- In the second row , corresponding to factor β_j *: All subscribers are present* in the CV σ_{ij}^2, so there are no coefficients to write (only r).
- In the same row corresponding to factor $\beta_{j,}$: The CV ra σ_j^2 its coefficient is ra , Why? Because in said CV, the subscribed i =1, 2,…a is not present, this level a being the one that is placed as a coefficient, in addition to r, that is: ra .
- In the third row corresponding to the interaction $(\alpha\beta)_{ij}$,: In the CV σ_{ij}^2 all subscribers are present, so there are no coefficients to write (only r).
- In the fourth row , for the experimental error variation factor, only its CV is written: σ_ε^2

Application of rule 4. In each row la ECM, define which CV stays and which disappears or is eliminated. This is perhaps the most important rule. Four cases are presented: It is explained based on the fact that all CVs are random; that is, the raw material is Table 5.

For any model in question, it is recommended to present it as if all its terms were random and then apply the corresponding rules.

a) α and β *random* ⇒Random model. It remains as in Table 5.
b) α and β, *fixed* ⇒Fixed, traditional, standard or classic model (the most used).
c) α *fixed* β random ⇒Mixed random model
d) α fixed ⇒ *random* βRandom mixed model

To exemplify these cases, we continue using model 1 as a pattern.

Illustration of case a : α_i , β_j random *s*⇒ Random model

It is common for these models to be referred to as class II or model II. For this case, based on model (1), Table 5 is derived, whose CV is explained below:

1. In all rows of the ECM, the following is written as the first CV: σ_ε^2
2. In the first row: *In* the CV r σ_{ij}^2 corresponding to the row of factor α_i , of the subscribers ij , I ignore i, (because we are in the row of factor α_i) , leaving j. Since j is random, the CV *r* σ_{ij}^2 , remains (it is not eliminated).
3. In the CV rb σ_i^2 of row α_i , whose subscriber is i, because it coincides exactly with the subscriber of factor α_i , the CV rb σ_i^2 remains. Whether fixed or random . This rule is considered as an alternative rule.
4. In the second row: In the CV, r σ_{ij}^2 *corresponding to the*

   factor row β_j ; of the subscribers ij , I ignore j (because we are in the queue of β_j ; leaving i). Since i of α_i , the CV r σ_{ij}^2 that remains is random.

5) In the CV ra σ_j^2 of the same row (of β_j), whose subscriber is j, because it exactly coincides with that of factor β_j , said CV remains. Whether fixed or random. This rule is considered an alternative rule.
6) In the third row, for the interaction: The CV r σ_{ij}^2 corresponding to the row of the interaction factor $(\alpha\beta)_{ij}$ whose subscribers are ij , because both coincide exactly; the CV remains. Whether fixed or random. This rule is also considered as an alternative rule.

Note in Table 5, since all FVs are random, all CVs remain.

Now, hypothesis testing for α_i, β_j (whose CVs σ_i^2, σ_j^2), can be clearly seen with the corresponding NDEs, in relation to la ECMthe interaction $(\alpha\beta)_{ij}$. The interaction can be done in relation to the NDE, which is clearly seen in the AdeV table, in Table 5.

Hypothesis test for factor α_i:

$$\text{Ho: } \sigma_i^2 = 0 \qquad VS \qquad \text{H1: } \sigma_i^2 \neq 0$$

$$Fc = \frac{(ECM)_{\alpha}}{(ECM)_{\alpha\beta}} = \frac{\sigma_\varepsilon^2 + rb\,\sigma_i^2 + r\,\sigma_{ij}^2}{\sigma_\varepsilon^2 + r\,\sigma_{ij}^2} = rb\,\sigma_i^2$$

Hypothesis test for factor β_j:

$$\text{Ho: } \sigma_j^2 = 0 \qquad VS \qquad \text{H}_1\text{: } \sigma_j^2 \neq 0$$

$$Fc = \frac{(ECM)_{\beta}}{(ECM)_{\alpha\beta}} = \frac{\sigma_\varepsilon^2 + ra\,\sigma_j^2 + r\,\sigma_{ij}^2}{\sigma_\varepsilon^2 + r\,\sigma_{ij}^2} = ra\,\sigma_j^2$$

Hypothesis test for the interaction factor $(\alpha\beta)_{ij}$:

$$\text{Hi: } \sigma_{ij}^2 = 0 \qquad VS \text{ H}_1\text{: } \sigma_{ij}^2 \neq 0$$

$$Fc = \frac{(ECM)_{\alpha\beta}}{(ECM)_{\varepsilon}} = \frac{\sigma_\varepsilon^2 + r\,\sigma_{ij}^2}{\sigma_\varepsilon^2} = r\,\sigma_{ij}^2$$

Consequently, for random models, tests are no longer done with the CM as if the effects were fixed: They are done using the NDE relationship, to know what to reject and what not to reject.

Illustration of case b : α_i fixed, β_j fixed, $\Rightarrow$Fixed model.

They usually point out these models as class 1 or model 1. Based on Table 5 and applying rule 4, Table 7 is derived. In all the rows, only two CVs appear. Why don't the rest appear? Let's see. :

1) In all rows of the ECM, it is written σ_ε^2 as the first CV.

2) For the first row corresponding to the row of the variation factor α_i, (Table 5): In the CV r σ^2_{ij},), of the subscribers ij , I ignore i (because we are in the row of the factor α_i , leaving j. As j is fixed, the CV r σ^2_{ij}, disappears.

3) In the same row corresponding to factor α_i , The CV rb σ^2_i whose subscriber is i, because it coincides exactly with that of factor α_i , the CV rb σ^2_i remains, whether fixed or random . This rule is considered an alternative rule.

4) In the second row corresponding to factor β_j , (table 1 *):* In the CV r σ^2_{ij} of subscribers ij , I ignore j, (because we are in the row of β_j), leaving i. Since i is fixed, the CV r σ^2_{ij} disappears.

5) In the same row corresponding to factor β_j ,: The CV ra σ^2_j whose subscriber j, by coinciding with that of factor β_j , remains, be it fixed or random . This rule is considered an alternative rule.

6) In the third row corresponding to the interaction $(\alpha\beta)_{ij}$: whose CV r , because both subscribers σ^2_{ij} ij coincide exactly , the CV r σ^2_{ij} remains, *sea* fixed or random . A This rule is considered an alternative rule. Consequently, Tables 6 and 7 are as follows:

Table 6. The AdeV of Model 2 CaA Design. Bifactorial.

Model		Fixed	Fixed	Int	Mistake	
		②	③	④	①	
Y_{ijk} =	μ+	α_i +	β_j +	$(\alpha\beta)_{ij}$ +	$\varepsilon_{k(ij)}$	(2)
		σ^2_i	σ^2_j	σ^2_{ij}	σ^2_ε	

Source: Own elaboration .

i = 1, 2, 3... a, levels of factor α_i *Fixed* .
j = 1, 2, 3... b, levels of factor β_j . *Fixed* .
k =1, 2, 3... r, repetitions of each combination of treatments (balanced case)

i, j, k: These are the symbolic subscripts.
a, b, r: These are the real subscripts.

Table 7. (from Table 5). For the Fixed Model

Dependent var.: Y					
ECM	↔	FV	GL	CM	ECM
① ② ④ $\sigma_\varepsilon^2 + rb\ \sigma_i^2 + r\sigma_{ij}^2$	↔	(fixed) Factor α_i	(a-1)	CM_α	① ② $\sigma_\varepsilon^2 + rb\sigma_i^2$
① ③ ④ $\sigma_\varepsilon^2 + ra\ \sigma_j^2 + r\sigma_{ij}^2$	↔ ↔	(Fixed) $_j$ factorβ	(b-1)	CM_β	① ③ $\sigma_\varepsilon^2 + ra\sigma_j^2$
① ④ $\sigma_\varepsilon^2 + r\sigma_{ij}^2$	↔ ↔	(Int.) Factor ($\alpha\beta$) $_{ij}$	(a-1)(b-1)	$CM_{\alpha\beta}$	① ④ $\sigma_\varepsilon^2 + r\sigma_{ij}^2$
① σ_ε^2	↔ ↔	Random. Error: $\varepsilon_{k(ij)}$	ab(r-1)	CM_ε	① σ_ε^2
		Total:	Apr-1	CM Total	

Source: Own elaboration

In the first column of Table 5, all the CVs that should be present when all terms are random, according to the established rules, are located. There, the rules are manipulated to define ECM in Table 7 (case b).

Important: It is common to use the letter indicating the name of the factor as σ_A^2 *subscripts :* , σ_B^2 , σ_{AB}^2 , instead of σ_i^2 , σ_j^2 , σ_{ij}^2 . To be consistent, it is advisable to use *symbolic subscripts* because the model is explained more clearly based on these *subscripts* , which is more than enough reason. Also, because many other things are done based on them:

1. The degrees of freedom (DF) column is defined in the AdeV table.
2. The expressions used to calculate the sum of squares for each variation factor are defined in the sum of squares column.
3. CVs are defined in the ECM column.
4. Sample size is defined.

It is important to note that tests for factors with all fixed effects models are done in the traditional way: Ratio of mean squares. In addition, interactions (significant) are more important than main effects, a relevant situation in factorial models.

Illustration of case c: α_i fixed, β_j , random ⇒Mixed Random Model.

From Table 4, model 1, and applying the rules already explained, Table 9 is derived. The presence or absence of CVs (whether they remain or not) is explained below:

1) In all rows of the ECM, it is written σ_ε^2 as the first CV.

2) In the first row of table 1: In the CV r σ^2_{ij}, in the row of factor α_i, of the subscribers ij , I ignore i, (because it is in the row of α_i), leaving j. As j is random, the CV r σ^2_{ij}, remains (it does not disappear).

3) In the same row as factor α_i ,: The CV rb σ^2_i, whose subscriber is i, because it coincides with that of factor α_i, the CV rb σ^2_i, remains either fixed or random.

4) Second row of factor β_j : In the CV r σ^2_{ij} of subscribers ij . I ignore j, because it is in the row of β_j, leaving i. Since i of α_i is fixed, the CV disappears.

5) In the same row as factor β_j : The CV ra σ^2_j, whose subscriber is j, because it coincides with that of factor β_j, remains, whether fixed or random.

6) Third row for the interaction $(\alpha\beta)_{ij}$: The CV σ^2_{ij}, whose subscribers are ij , because both coincide, the CV σ^2_{ij}, remains (whether fixed or random).

Table 8. The AdeV of Model 3
CaA design. Bifactorial.

		Fixed	Random.	Int (Random)	Mistake	
Model		②	③	④	①	
Y_{ijk} =	μ+	α_i + σ_i^2	β_j + σ_j^2	$(\alpha\beta)_{ij}$ + σ_{ij}^2	$\varepsilon_{k(ij)}$ σ_ε^2	(3)

Source: Own elaboration.

i = 1, 2, 3...at levels of factor α_i , *Fixed.*
J = 1,.2, 3...b levels of factor β_j , *Random*
J = 1,.2, 3...r repetitions (balanced case)

i, j, k, are the *symbolic subscripts.*
a, b, r, are the *real subscripts*

Table 9 . (from Table 5): For the Mixed Random Model

Dependent var: Y					
ECM	↔	FV	GL	CM	ECM
① ② ④ $\sigma_\varepsilon^2 + rb\ \sigma_i^2 + r\sigma_{ij}^2$	↔ ↔	(fixed) Factor α_i	(a-1)	CM_α	① ② ④ $\sigma_\varepsilon^2 + rb\ \sigma_i^2 + r\sigma_{ij}^2$
① ③ ④ $\sigma_\varepsilon^2 + ra\ \sigma_j^2 + r\sigma_{ij}^2$	↔ ↔	(Random) Factor β_j	(b-1)	CM_β	① ③ $\sigma_\varepsilon^2 + ra\,\sigma_j^2$
① ④ $\sigma_\varepsilon^2 + r\sigma_{ij}^2$	↔ ↔	Int (Random) Factor $(\alpha\beta)_{ij}$	(a-1)(b-1)	$CM_{\alpha\beta}$	① ④ $\sigma_\varepsilon^2 + r\sigma_{ij}^2$
① σ_ε^2	↔ ↔	(Random) Error: $\varepsilon_{k(ij)}$	ab(r-1)	CM_ε	① σ_ε^2
		Total:	Abr-1	CM_{To}	

Source: Own elaboration

The first column corresponds to Table 4, which is where the rules are manipulated to define the last column, corresponding to Table 9; with practice, this is no longer necessary. You work directly with the terms of the model.

Significance tests: Note in Table 9: The significance tests for the $\beta_{j\ factor}$ and Interaction $(\alpha\beta)_{ij}$, are done in the traditional way like this:

$$Fc\ .\ = (CM)_\beta/(CM)_\varepsilon \quad \text{and} \quad Fc\ .\ = (CM)_{\alpha\beta}/(CM)_\varepsilon$$

However, this is not the case for the significance test for the variation factor α_i, which will have to be determined with the following ECM relation, where (ECM) $_{\alpha\beta}$is taken as the error term.

Ho: $\sigma_i^2 = 0$ *VS* H1: $\sigma_i^2 \neq 0$

$$Fc = \frac{(ECM)_{\alpha}}{(ECM)_{\alpha\beta}} = \frac{\sigma_\varepsilon^2 + rb\,\sigma_i^2 + r\,\sigma_{ij}^2}{\sigma_\varepsilon^2 + r\,\sigma_{ij}^2} = rb\,\sigma_i^2$$

Case illustration d : α_i random β_j fixed: Mixed model Random .

Opposite case to the above. How are the ECMs in each of the rows? Following the same reasoning in Table , Tables 10 and 11 are generated:

Table 10 . The AdeV of Model 4 Completely randomized design (CaA). Bifactorial.

	Alat .	Fixed	Int (Random)	Mistake	
Model	②	③	④	①	
$Y_{ijk} =$	$\mu +$ $\alpha_{Yo} +$	β_j +	$(\alpha\beta)_{ij} +$	$\varepsilon_{k(ij)}$	(4)
	σ_i^2	σ_j^2	σ_{ij}^2	σ_ε^2	

Source: Own elaborati

i = 1, 2, 3... a, levels of factor α_i . *Random*
j = 1, 2, 3... b, levels of factor β_j *Fixed*
k =1, 2, 3... r, repetitions (Balanced case)

i, j, k: *Symbolic subscripts.*

a, b, r: Real subscripts.

Table 11 . (from Table 5). For the Random Mixed Model.

Dependent Var.: Y					
ECM	↔	FV	GL	CM	ECM
① ② ④ $\sigma_\varepsilon^2 + rb\,\sigma_i^2 + r\,\sigma_{ij}^2$	↔ ↔	(Random) Factor α_i	(a-1)	CMα	① ② $\sigma_\varepsilon^2 + rb\,\sigma_i^2$
① ③ ④ $\sigma_\varepsilon^2 + ra\,\sigma_j^2 + r\,\sigma_{ij}^2$	↔ ↔	(Fixed) Factor β_j	(b-1)	CMβ	① ③ ④ $\sigma_\varepsilon^2 + ra\,\sigma_j^2 + r\,\sigma_{ij}^2$
① ④ $\sigma_\varepsilon^2 + r\,\sigma_{ij}^2$	↔ ↔	Int (random) Factor 8αβ)$_{ij}$	(a-1)(b-1)	CMαβ	① ④ $\sigma_\varepsilon^2 + r\,\sigma_{ij}^2$
① σ_ε^2	↔	Error: $\varepsilon_{k(ij)}$	ab(r-1)	CMε	① σ_ε^2
		Total:	Abr-1	CM To	

Source: Own elaboration.

Significance tests for factors α_i and int . $(\alpha\beta)_{ij}$, in relation to the CV σ_ε^2 in traditional form, as follows: Fc = $(CM)_\beta$/ $(CM)_\varepsilon$and Fc = $(CM)_{\alpha\beta}$/ $(CM)_\varepsilon$

However, this is not the case for the variation factor β_j, which will have to be done with the following ECM relation, where $(ECM)_{\alpha\beta}$is taken as the error term:

Hi: $\sigma_j^2 = 0$ *VS* H1: $\sigma_j^2 \neq 0$

$$Fc = \frac{(ECM)_\beta}{(ECM)_{\alpha\beta}} = \frac{\sigma_\varepsilon^2 + ra\,\sigma_j^2 + r\,\sigma_{ij}^2}{\sigma_\varepsilon^2 + r\,\sigma_{ij}^2} = ra\,\sigma_j^2$$

More complicated models. Calculation of NDEs

A first example is shown: suppose the following three-factor model (α, β, γ, all random with 4, 2 and 5 levels). Structure: factorial housed in a BaA with 3 blocks, balanced case (see Table 13, model 5).

i = 1, 2…a = 4 levels of factor α_i : *random*
j = 1, 2…b = 2 levels of factor β_j : *random*
k = 1, 2....c = 5 levels of factor γ_k : *random*
l = 1, 2....r = 3 repetitions ρ_l : *random*

i, j, k, l: *symbolic subscripts.*

a, b, c, r: *real subscripts.*

Table 12 . The AdeV of Model 5

Model		TO ②	TO ③	TO ④	TO ⑤	TO ⑥	TO ⑦	TO ⑧	TO ⑨	TO ①	
And $_{ijk}$ =	μ+	ρ_l +	α_i +	β_j +	$(\alpha\beta)_{ij}$ +	γ_k +	$(\alpha\gamma)_{ik}$ +	$(\beta\gamma)_{jk}$ +	$(\alpha\beta\gamma)_{ijk}$ +	$\varepsilon_{k(ij)}$	(5)
			σ_i^2	σ_j^2	σ_{ij}^2	σ_k^2	σ_{ik}^2	σ_{jk}^2	σ_{ijk}^2	σ_ε^2	
A ⇒Random With $\varepsilon_{k(ij)} \sim NI(0, \sigma_\varepsilon^2$											

Source: Own elaboration

Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

Table 13. For the random three-factor model.

Variable	dependent:	Y	
FV	GL	ECM: In this column the CVs are selected, according to the type of factors	ECM: In this column the coefficients are replaced by their actual values.
Block .	(r-1)=2	The block test is not valid (because designs using blocks are not randomized). Testing blocks is not scientifically interesting.	
Treat	(t-1)=39	No NDEs are reported for treatments, because they are already considered in the breakdown of their main effects and interactions.	
Factor α_I	(a-1)=3	① ③ ⑤ ⑦ ⑨ $\sigma_\varepsilon^2 + \text{rbc}\ \sigma_i^2 + \text{rc}\ \sigma_{ij}^2 + \text{rb}\ \sigma_{ik}^2 + \text{r}\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 30\ \sigma_i^2 + 15\ \sigma_{ij}^2 + 6\ \sigma_{ik}^2 + 3\ \sigma_{ijk}^2$
Factor β_j	(b-1) =1	① ④ ⑤ ⑧ ⑨ $\sigma_\varepsilon^2 + \text{rac}\ \sigma_j^2 + \text{rc}\ \sigma_{ij}^2 + \text{ra}\ \sigma_{jk}^2 + \text{r}\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 60\ \sigma_j^2 + 15\ \sigma_{ij}^2 + 12\ \sigma_{jk}^2 + 3\sigma_{ijk}^2$
Int: $(\alpha\beta)_{ij}$	(a-1)(b-1)=3	① ⑤ ⑨ $\sigma_\varepsilon^2 + \text{rc}\ \sigma_{ij}^2 + \text{r}\sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 15\ \sigma_{ij}^2 + 3\sigma_{ijk}^2$
Factor: γ_k	(c-1)=4	① ⑥ ⑦ ⑧ ⑨ $\sigma_\varepsilon^2 + \text{rab}\ \sigma_k^2 + \text{rb}\ \sigma_{ik}^2 + \text{ra}\ \sigma_{jk}^2 + \text{r}\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 24\ \sigma_k^2 + 6\ \sigma_{ik}^2 + 12\ \sigma_{jk}^2 + 3\ \sigma_{ijk}^2$
Int: $(\alpha\gamma)_{ik}$	(a-1)(c-1)=12	① ⑦ ⑨ $\sigma_\varepsilon^2 + \text{rb}\ \sigma_{ik}^2 + \text{r}\sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 6\ \sigma_{ik}^2 + 3\sigma_{ijk}^2$
Int: $(\beta\gamma)_{jk}$	(b-1)(c-1)=4	① ⑧ ⑨ $\sigma_\varepsilon^2 + \text{ra}\ \sigma_{jk}^2 + \text{r}\sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 12\ \sigma_{jk}^2 + 3\sigma_{ijk}^2$
Int: $(\alpha\beta\gamma)_{ijk}$	(a-1) (b-1) (c-1) =12	① ⑨ $\sigma_\varepsilon^2 + \text{r}\sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 3\sigma_{ijk}^2$
Mistake: $\varepsilon_{k(ij)}$	(r-1)(abc-1)=78	① σ_ε^2	σ_ε^2
Total: Y_{ijk}			

Source: Own elaboration.

Since all principal factors are random, so are the interactions. However, it is worth remembering that the variation factors can be fixed or random. Note that the corresponding notation is located in the following parts:

a) Above and below each term of the corresponding model.

b) In the part where the symbolic and real levels are defined.

c) In the FV column, in the AdeV table.

With this practice, the researcher in training remembers or constantly keeps in mind which factors are fixed and which factors are random, which greatly facilitates the

management of the subscripts to determine which CVs remain or do not remain in the ECM column, to know which hypothesis to reject and which hypothesis not to reject.

Both in the model and of course in the AdeV table, the term blocks appears (because the model by construction includes the term blocks). When the characteristics of the research do not require the use of blocks, this means that the researcher is almost certainly using a CaA design.

In the present example, since all factors are random, they all remain, they do not disappear. Significance tests for the FVs that can or cannot be done are determined with an appropriate NDE relationship.

Significance test for variation factor α_i :

Ho: $\sigma_i^2 = 0$ *VS* H1: $\sigma_i^2 \neq 0$

$$Fc = \frac{(ECM)_{\alpha} + (ECM)_{\alpha\beta\gamma}}{(ECM)_{\alpha\beta} + (ECM)_{\alpha\gamma}} = \frac{\overset{①}{\sigma_\varepsilon^2} + \overset{③}{rbc}\,\overset{⑤}{\sigma_i^2} + \overset{⑦}{rc}\,\overset{⑨}{\sigma_{ij}^2} + \overset{①}{rb}\,\overset{⑦}{\sigma_{ik}^2} + \overset{⑨}{r}\,\overset{①}{\sigma_{ijk}^2} + \overset{⑨}{\sigma_\varepsilon^2} + r\sigma_{ijk}^2}{\sigma_\varepsilon^2 + rc\,\sigma_{ij}^2 + r\,\sigma_{ijk}^2 + \sigma_\varepsilon^2 + rb\,\sigma_{ik}^2 + r\,\sigma_{ijk}^2} = rb\,\sigma_\varepsilon^2$$

$$Fc = \frac{(ECM)_{\alpha} + (ECM)_{\alpha\beta\gamma}}{(ECM)_{\alpha\beta} + (ECM)_{\alpha\gamma}} = \frac{\overset{①}{\sigma_\varepsilon^2} + \overset{③}{rbc}\,\overset{⑤}{\sigma_i^2} + \overset{⑦}{rc}\,\overset{⑨}{\sigma_{ij}^2} + \overset{①}{rb}\,\overset{⑦}{\sigma_{ik}^2} + \overset{⑨}{r}\,\overset{①}{\sigma_{ijk}^2} + \overset{⑨}{\sigma_\varepsilon^2} + r\sigma_{ijk}^2}{\sigma_\varepsilon^2 + rc\,\sigma_{ij}^2 + rb\,\sigma_{ik}^2 + \sigma_{ijk}^2 + \sigma_\varepsilon^2 + r\,\sigma_{ijk}^2} = rb\,\sigma_\varepsilon^2$$

$$Fc = \frac{(ECM)_{\alpha} + (ECM)_{\alpha\beta\gamma}}{(ECM)_{\alpha\beta} + (ECM)_{\alpha\gamma}} = \frac{\overset{①}{\sigma_\varepsilon^2} + \overset{③}{30}\,\overset{⑤}{\sigma_i^2} + \overset{⑦}{15}\,\overset{⑨}{\sigma_{ij}^2} + \overset{①}{6}\,\overset{⑦}{\sigma_{ik}^2} + \overset{⑨}{3}\,\overset{①}{\sigma_{ijk}^2} + \overset{⑨}{\sigma_\varepsilon^2} + 3\sigma_{ijk}^2}{\sigma_\varepsilon^2 + 15\,\sigma_{ij}^2 + 6\,\sigma_{ik}^2 + \sigma_{ijk}^2 + \sigma_\varepsilon^2 + 3\,\sigma_{ijk}^2} = 30\,\sigma_i^2$$

$$\therefore \quad Fc = 30\,\sigma_i^2$$

With the above, the following was done (didactically):

1. In the first significance test, the CV coefficients were symbolically placed in the numerator and denominator.

2. In the second, the CVs are rearranged in such a way that they correspond: each of the CVs of the numerator with each of the CVs of the denominator, leaving the term to be tested alone or isolated: 30 $\sigma_i^2 = 30\,\sigma_i^2$
3. In the third, the coefficients have been replaced by their real values, leaving the term to be tested alone or isolated: 30 σ_i^2 The simplification is precisely 30 σ_i^2. The value Fc , is obtained by multiplying 30 by the value of the variance σ_i^2 .

On the other hand: F ~F (3+12, 3+12 and α) = F (15, 15, 0.5) = 2.41

Finally, the F value of the tables (2.41) is compared with the value of Fc (30 σ_i^2).

Significance test for the variation factor γ_k :

$$H_o: \sigma_\gamma^2 = 0 \qquad VS \qquad H_1: \sigma_\gamma^2 \neq 0$$

$$Fc = \frac{(ECM)_\gamma + (ECM)_{\alpha\beta\gamma}}{(ECM)_{\alpha\gamma} + (ECM)_{\beta\gamma}} = \frac{\overset{①\ ⑥\ \ ⑦\ \ \ ⑧\ \ \ ⑨\ \ ①\ \ ⑨}{\sigma_\varepsilon^2 + rab\ \sigma_k^2 + rb\ \sigma_{ik}^2 + ra\ \sigma_{jk}^2 + r\ \sigma_{ijk}^2 + \sigma_\varepsilon^2 + r\sigma_{ijk}^2}}{\underset{①\ \ ⑦\ \ \ ⑨\ \ \ ①\ \ ⑧\ \ \ \ ⑨}{\sigma_\varepsilon^2 + rb\ \sigma_{ik}^2 + r\ \sigma_{ijk}^2 + \sigma_\varepsilon^2 + ra\ \sigma_{jk}^2 + r\sigma_{ijk}^2}} = 24\,\sigma_k^2$$

As in the previous case: Fc . ~ F_{tables} (4+12, 12+4, α=F (16, 16, 0.05) = 2.33

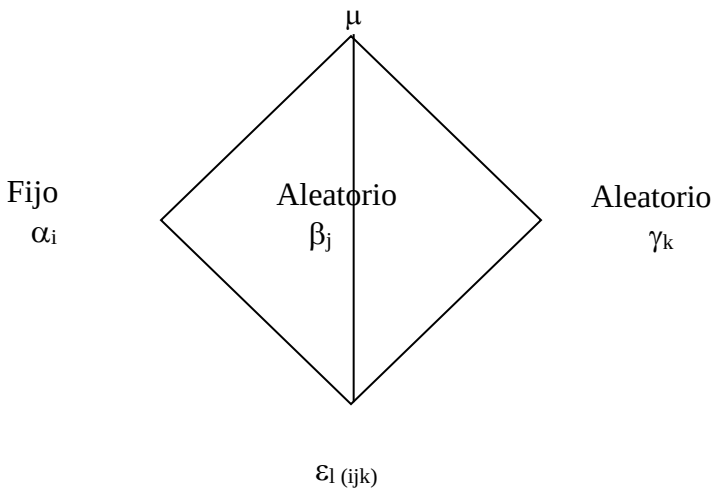
Finally, the F value of tables (2.33) is compared with the Fc value

Second example: Suppose (as in example 1, the following three-factor model: α_i , β_j , γ_k with 4, 2 and 5 levels respectively, in a BaA design, with 3 blocks; but now, with a fixed factor α_i and two random factors β_j and γ_k .

i = 1, 2…a = 4 levels of factor α_i : *(fixed)*
j = 1, 2…b = 2 levels of factor β_j : *(random)*
k = 1, 2....c = 5 levels of factor γ_k : *(random)*
l = 1, 2….r = 3 repetitions. ρ_i : *(random)*

i, j, k, l: *symbolic subscripts.*
a, b, c, r: *real subscripts*

Figure 2. Representation of the three-factor model.



Source: Own elaboration.

Table 14. The AdeV of Model 6

			F	TO	TO	TO	TO	TO	TO	TO	
Model		②	③	④	⑤	⑥	⑦	⑧	⑨	①	
Y_{ijk} =	$\mu+$	ρ_l+	α_i+	β_j+	$(\alpha\beta)_{ij}+$	γ_{k+}	$(\alpha\gamma)_{ik}+$	$(\beta\gamma)_{jk}+$	$(\alpha\beta\gamma)_{ijk}+$	$\varepsilon_{k(ij)}$	(6)
			σ_i^2	σ_j^2	σ_{ij}^2	σ_k^2	σ_{ik}^2	σ_{jk}^2	σ_{ijk}^2	σ_ε^2	

Source: Own elaboration

Revista Iberoamericana para la Investigación y el Desarrollo Educativo
ISSN 2007 - 7467

Table 15. For the fixed three-factor model and two random ones.

Variable	Dependent:	Y	
FV	GL	ECM: In this column the CVs are selected, according to the type of factors	ECM: In this column the coefficients are replaced by their actual values.
Blocks: ρ_j	(r-1) = 2	The block test is not valid (because designs that use blocks are not randomized). Testing blocks is not scientifically interesting .	
Treat	(t-1) = 39	No NDEs are reported for treatments, because they are already considered in the breakdown of their main effects and interactions.	
Factor: α_j	(a-1) = 3	① ③ ⑤ ⑦ ⑨ $\sigma_\varepsilon^2 + rbc\ \sigma_i^2 + rc\ \sigma_{ij}^2 + rb\ \sigma_{ik}^2 + r\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 30\ \sigma_i^2 + 15\ \sigma_{ij}^2 + 6\ \sigma_{ik}^2 + 3\ \sigma_{ijk}^2$
Factor: β_j	(b-1) = 1	① ④ ⑧ ⑨ $\sigma_\varepsilon^2 + rac\ \sigma_j^2 + ra\ \sigma_{jk}^2 + r\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 60\ \sigma_j^2 + 12\ \sigma_{jk}^2 + 3\ \sigma_{ijk}^2$
Int: $(\alpha\beta)_{ij}$	(a-1) (b-1) = 3	① ⑤ ⑨ $\sigma_\varepsilon^2 + rc\ \sigma_{ij}^2 + r\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 15\ \sigma_{ij}^2 + 3\ \sigma_{ijk}^2$
Factor: γ_k	(c-1) = 4	① ⑥ ⑧ $\sigma_\varepsilon^2 + rab\ \sigma_k^2 + ra\ \sigma_{jk}^2$	$\sigma_\varepsilon^2 + 24\ \sigma_k^2 + 12\ \sigma_{jk}^2$
Int: $(\alpha\gamma)_{ik}$	(a-1) (c-1)=12	① ⑦ ⑨ $\sigma_\varepsilon^2 + rb\ \sigma_{ik}^2 + r\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 6\ \sigma_{ik}^2 + 3\ \sigma_{ijk}^2$
Int: $(\gamma)_{jk}$	(b-1) (c-1) = 4	① ⑧ $\sigma_\varepsilon^2 + ra\ \sigma_{jk}^2$	$\sigma_\varepsilon^2 + 12\ \sigma_{jk}^2$
Int: $(\alpha\beta\gamma)_{ijk}$	(a-1)(b-1)(c-1) =12	① ⑨ $\sigma_\varepsilon^2 + r\ \sigma_{ijk}^2$	$\sigma_\varepsilon^2 + 3\ \sigma_{ijk}^2$
Error: $\varepsilon_{k(ij)}$	(r-1)(abc-1)=78	① σ_ε^2	σ_ε^2
Total: Y_{ijk}	rabc-1=119		

Source: Own elaboration.

The hypothesis test for CV σ_i^2, variation factor α_i is expressed as follows:

$$H_o: \sigma_i^2 = 0 \qquad VS \qquad H_1: \sigma_i^2 \neq 0$$

$$Fc = \frac{(ECM)_\alpha + (ECM)_{\alpha\beta\gamma}}{(ECM)_{\alpha\beta} + (ECM)_{\alpha\gamma}} = 30\,\sigma_i^2$$

Similar to the previous case, you can check it.

Hypothesis test for factor γ_k is expressed as follows:

$$Fc = \frac{(ECM)_{\gamma}}{(ECM)_{\beta\gamma}} = \frac{\sigma_{\varepsilon}^2 + rab\ \sigma_k^2 + ra\,\sigma_{jk}^2}{\sigma_{\varepsilon}^2 + ra\,\sigma_{jk}^2} = rab\ \sigma_k^2$$

$$Fc = \frac{(ECM)_{\gamma}}{(ECM)_{\beta\gamma}} = \frac{\sigma_{\varepsilon}^2 + 24\ \sigma_k^2 + 12\,\sigma_{jk}^2}{\sigma_{\varepsilon}^2 + 12\,\sigma_{jk}^2} = 24\,\sigma_k^2$$

$$\sigma_k^2 \sim F(4, 4, 0.05) = 6.39$$

The approximate test is performed by comparing Fc , with the f value from tables (6.39)

Third example *:* With a mixed (random) four-factor model. Three crossed factors: α_i , β_j , δ_l and a fourth: γ_k : hierarchical (nested in α_i , β_j), with 3 blocks. Of the four factors we will assume fixed (F): α_i , δ_l and random (A): β_j and γ_k .

Table 16. The AdeV of Model 7

		F	A	A	A	F	F	A	A	A	A	
Mod	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	11	①	(7)
$Y_{ijk} =$	$\mu +$	$\rho_m +$	$\alpha_i +$	$\beta_j +$	$(\alpha\beta)_{ij} +$	$\gamma_{k(ij)} +$	$\delta_l +$	$(\alpha\delta)_{il} +$	$(\beta\delta)_{jl} +$	$(\alpha\beta\delta)_{ijl} +$	$(\gamma\delta)_{k(ij)l} +$	$\varepsilon_{m(ijkl)}$	
		σ_i^2	σ_j^2	σ_{ij}^2	$\sigma_{k(ij)}^2$	σ_l^2	σ_{il}^2	σ_{jl}^2	σ_{ijl}^2	$\sigma_{k(ij)l}^2$	$\sigma_{m,(ijkl)}^2$	

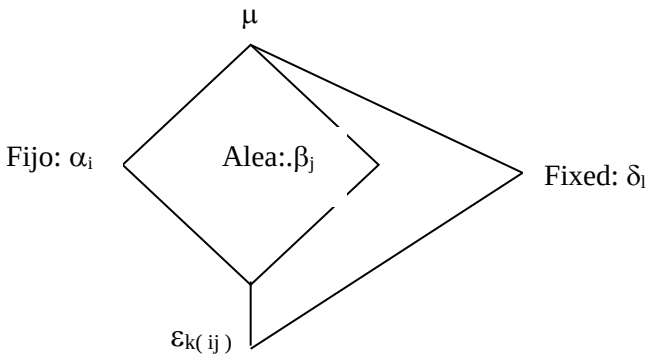
Source: Own elaboration

i = 1, 2… a⇒ α_i : Fixed and crossed with β_j and δ_l
j = 1, 2… b⇒ β_j : Random and crossed with α_i and δ_l
k = 1, 2.... c⇒ $\gamma_{k(ij)}$: Random and nested in α_i , β_j
l = 1, 2 …d⇒ δ_l : Fixed and crossed with α_i , β_j , $\gamma_{k(ij)}$
m = 1, 2... r⇒ ρ_m : Repetitions

i, j, k, l, m: *symbolic subscripts.*

a, b, c, r: *real subscripts*

Figure 3. Representation of the four-factor model.



Source: Own elaboration

Table 17. For the random mixed four-factor model .

Variable Dependent: Y		
FV	GL	ECM
Block	r-1= 2	Invalid test
Random. Factor α_i	a-1 = 3	① ③ ⑤ ⑥ $\sigma_\varepsilon^2 + \text{bcdr}\ \sigma_i^2 + \text{cdr}\ \sigma_{ij}^2 + \text{dr}\ \sigma_{k(ij)}^2$
Random. Factor β_j	b-1 = 1	① ④ ⑥ $\sigma_\varepsilon^2 + \text{acdr}\ \sigma_j^2 + \text{dr}\ \sigma_{k(ij)}^2$
Random. Int. $(\alpha\beta)_{ij}$	(a-1) (b-1) = 3	① ⑤ ⑥ $\sigma_\varepsilon^2 + \text{cdr}\ \sigma_{ij}^2 + \text{dr}\ \sigma_{k(ij)}^2$
Random. $\gamma_{k(ij)}$	(c-1) ab = 32	① ⑥ $\sigma_\varepsilon^2 + \text{dr}\ \sigma_{k(ij)}^2$
Fixed δ_l	d-1 = 2	① ⑦ ⑨ $\sigma_\varepsilon^2 + \text{abcr}\ \sigma_l^2 + \text{bcr}\ \sigma_{jl}^2$
Fixed $(\alpha\delta)$ the	(a-1(d-1) = 6	① ⑧ ⑩(11) $\sigma_\varepsilon^2 + \text{bcr}\ \sigma_{il}^2 + \text{cr}\ \sigma_{ijl}^2 + \text{r}\ \sigma_{k(ij)l}^2$
Random. $(\beta\delta)_{jl}$	(b-1) (d-1) = 2	① ⑨(11) $\sigma_\varepsilon^2 + \text{bcr}\ \sigma_{jl}^2 + \text{r}\ \sigma_{k(ij)l}^2$
Random. $(\alpha\beta\delta)_{ijl}$	(a-1) (b-1) (d-1) = 6	① ⑩(11) $\sigma_\varepsilon^2 + \text{cr}\ \sigma_{ijl}^2 + \text{r}\ \sigma_{k(ij)l}^2$
Random. $(\gamma\delta)_{k(ij)l}$	(c-1) ab(d-1) = 64	①11 $\sigma_\varepsilon^2 + \text{r}\ \sigma_{k(ij)l}^2$
Mistake	By diff . = 226	① σ_ε^2
Total	Ijklm-1 = 359	

Source: Own elaboration

In Model 7 and the AdeV table, an EdeR should have been included after the term blocks. This was not done in order to pay more attention to the construction of the *Expected Values* (EV) in the ECM column. The interested reader can do it as an exercise to verify that the hypothesis test on blocks is not possible, and that it is not of scientific interest.

In model (7), some *interactions* were not included and certainly not in the AdeV table, for example: $(\alpha\gamma)_{k(ij)}$, because *γit is a nested factor in* α_i as is the term ⑥: $\gamma_{k(ij)}$ and also because the combinations of subscribers of these terms already exist in said term. For the same reason, the interactions *:* $(\beta\gamma)_{jk}$, $(\alpha\beta\gamma)_{k(ij)}$, $(\alpha\gamma\delta)_{k(ij)l}$, $(\beta\gamma\delta)_{k(ij)l}$, $(\alpha\beta\gamma)_{k(ij)l}$, are not included, which is seen more clearly in the graph. The procedure to determine which CVs remain (according to rule 4) is explained below, in particular for the *nested terms* .

So far, reference has been made to *fixed, random and mixed factors and to balanced complete designs and models.* In model (7), Table 17 and in general in many models that are presented in real life according to the phenomenon being studied, to estimate CV, two other very interesting cases are presented :

Case 1: Nested factors (also called hierarchical)

Case 2: Cross factors.

Nested factors: For example $\gamma_{k(ij)}$. This means that the k levels are nested within i, j. Each level of k is combined with a single level of i and with a single level of j. In balanced nested models , all the levels of the nested factor correspond to the levels of the factor in which it is nested.

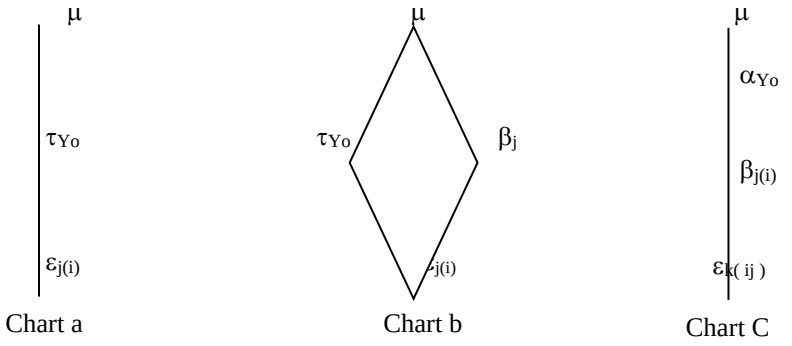
Nested factors are expressed using parentheses . Subscripts outside the parentheses are those nested within subscripts inside the parentheses. For example:

a) In the completely randomized design model, the random error $\varepsilon_{j(i)}$ is nested in the treatments factor.

b) In a model with two study factors without interaction, the random error is nested in both terms .

c) In the model with two study or variation factors, one of which (any one) is nested in the other, it is written like this: $\beta_{j(i)}$ and the random error is nested in both: $\varepsilon_{k(ij)}$

Figure 4. Nested factors



Source: Own elaboration

And so on. However, in practice, errors are usually written like this; ε_{ij} , ε_{ijk} , which does not indicate that the errors are nested. With nested factors there is no interaction.

On the other hand, crossed factors occur in combination with each level of another factor, that is, when each level of a factor is tested at each level of another factor. It has the characteristic of crossing a factor with levels of another factor; unlike nested factors, the levels of all the factors cross each other. Remember that the factors are independent variables.

Every researcher should keep in mind that crossed factors are those in which each of the subscribers of one factor is combined with each of the subscribers of the other factor (all with all). In crossed models, the number of repetitions must be the same for each of the combinations of levels (the treatments).

Discussion

Statistical models are a set of procedures that allow us to learn from quantitative data in a reliable manner and draw conclusions about them in a reasonable or doubtful manner. It is a mathematical relationship between random and non-random variables.

Furthermore, it is important to identify that in complete models all research units (experimental units), (investigations via experiment), the same number is maintained until the end of the research. And, that balanced models are those in which all treatments or combinations of treatments are repeated the same number of times.

It is emphasized that if, due to fortuitous circumstances, the model in question is not complete and balanced, it is an incomplete and unbalanced model. To avoid this situation,

the researcher must take extreme care in planning and managing or conducting the research project.

Incomplete and unbalanced models are frequently found in observational studies or projects. Models of this type can occur consistently in social sciences such as sociology, psychology, anthropology, economics, and also in medicine, mainly in retrospective research studies or projects.

Observational models have no experimental basis and can be managed with regression models. The experimental statistical design corresponds to a linear model that contains effects and a random error nested within the rest of the effects.

More than a research article, it is a didactic text for improving the use and application of ECM, which is rarely written about. As Restrepo (2007) rightly says, "When carrying out a variance analysis, the type of factor or factors involved in the experimental classification design must be taken into account, in order to generate the appropriate ECM, and thus reach coherent conclusions in the analysis of the information" (p. 201) .

By way of conclusion

Some recommendations and rules to obtain the ECM, in relation to rule 4, which determines the CVs that remain and those that do not remain in the corresponding row, in relation to the nested factors, it is emphasized that it is *important* to start from the model, manipulating the corresponding symbolic and real subscribers or subscripts. Likewise, it is *convenient to list* the terms of the model. *For each Source of Variation,* select the CVs that contain at least one subscriber equal to said factor; *determine* by manipulating subscribers , the coefficients of each selected CV, one of which is always r, except for σ_{ε}^{2} and *determine the subscribers that remain and those that disappear* how? Of two or more subscribers, the subscriber(s) equal to that of the factor in question *is ignored* . Of the subscribers that remain, *if they are fixed, the CV disappears . If they are random, the CV remains, does not disappear* . Furthermore, by the alternative rule, the CV whose subscribers exactly match those of the FV remains in the queue (whether fixed or random).

If any factor is nested in one or more factors (Example 4), proceed as *upon* in table 17 of the AdeV, model 7, which is in the Variation Factor $\gamma_{k(ij)}$, and that we are analyzing term 11 of the model: $(\gamma\delta)_{k(ij)l}$, whose CV is $\sigma^{2}_{k(ij)l}$ *to see if it stays or not* . We proceed by ignoring the subscribers ij (inside the parentheses) it is considered as $\sigma^{2}_{kl\,k}$ being a nested factor , γthe

following is said: as we are in row k, of the subscribers kl, I ignore k and we are left with 1. As 1 is fixed, CV 11 disappears from the row considered.

It is also important *to assume* in the AdeV table that we are in the FV : $(\gamma\delta)_{k(ij)l}$ and that we are analyzing term 11 of the model $(\gamma\delta)_{k(ij)l}$, to see if its CV, $\sigma^2_{k(ij)l}$, *stays or does not stay* . It is considered again, as σ^2_{kl}; that is, the subscribers i, j (inside the parentheses) are ignored, and we say the following: since we are in row l (ele), of the subscribers kl, I ignore 1 and we are left with k. Since k is random, the CV 11: $\sigma^2_{k(ij)l}$ stays, does not disappear. Furthermore, by alternative rule, since the subscribers considered coincide, the CV $\sigma^2_{k(ij)l}$, does not disappear (whether fixed or random).

For the FV $(\alpha\delta)_{il}$ (the term ⑧of the model), whose CV is σ^2_{il}, because of the coincidence of subscribers (by the alternative rule) said CV remains. The terms ⊃and 11 remain because j and k, are random. *The terms* ⑨and ⑩the model: $(\beta\delta)_{jl}$ and $(\alpha\beta\delta)_{ijl}$ respectively, is the same case of the FV $(\alpha\delta)_{il}$ (model term) and so on.∩

Future lines of research

Using the experimental design chosen by the researcher leads to obtaining the ECM, which is also necessary in the analysis of variance. The ECM is a topic that can be learned in statistics or quantitative research courses, but it should be noted that there is little literature on the subject, especially in Spanish; therefore, writing and socializing on the subject represents an opportunity to enhance and facilitate its use in support of experimental researchers who seek to learn about the ECM to determine statistical tests and be able to decide whether to accept or reject the hypothesis of interest.

References

Cienfuegos Ibarra, F. (Comunicación personal, 2 mayo de 1990). Notas del curso Experimentación Agrícola. Facultad de Estudios Superiores Cuautitlán (FESC-C), Universidad Nacional Autónoma de México (UNAM).

Fisher, R. A. (1936). El uso de múltiples mediciones en problemas taxonómicos. *Anales de Eugenesia*, 7, 179-188. http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x

Martínez Garza A. (1988). *Diseños Experimentales. Métodos y elementos de Teoría*. Edit. Trillas.

Méndez Ramírez I. (1984). El Protocolo de Investigación. Lineamientos para su elaboración y análisis. Edit. Trillas.

Restrepo, L. F. (2007). La esperanza del cuadrado medio. Revista Colombiana De Ciencias Pecuarias, 20(2), 9. https://doi.org/10.17533/udea.rccp.324136

Santizo Rincón A. (1974). Curso Experimentación Agrícola. Centro de Estadística y Cálculo (CEC). Colegio de Postgraduados (CP), Chapingo, Méx.