International Journal of Scientific Research Updates
Journal homepage: https://orionjournals.com/ijsru/
ISSN: 2783-0160 (Online)

(Research Article)

The generation of the experimental design

João Gilberto Corrêa da Silva *

Department of Mathematics and Statistics, Federal University of Pelotas, Pelotas, RS, Brazil.

International Journal of Scientific Research Updates, 2022, 03(02), 064–069

Publication history: Received on 12 April 2022; revised on 16 May 2022; accepted on 19 May 2022

Article DOI: https://doi.org/10.53430/ijsru.2022.3.2.0043

Abstract

Texts and teaching of Experimental Statistics emphasize the statistical analysis of experiments and make only references to the conceptual foundation of experimental research. Basic concepts are vaguely and incompletely defined, and experimental designs are presented as a set of recipes from which one must be chosen for each particular experiment. Consequently, inferences derived from the experiment are often biased. A conceptual basis for the experimental research is considered and a rational procedure for generating the experimental design is suggested. The generation of the experimental design is based on the separate definitions of the structures of the experimental factors and the unit factors, and the association of these two structures by randomization and presence of the experimental factors in the sample. This approach leads to the clear identification of the confounding of the effects from these two structures and of the experimental errors that affect the effects of the experimental factors.

Keywords: Experiment; Experimental Research; Experimental Factor; Treatment Factor; Intrinsic Factor; Condition Structure; Unit Structure; Randomization; Experiment Structure

1 Introduction

Fisher observes that the design of the experiment can be considered as the association of two independent structures: one related to the questions of interest to be answered by the experiment (*structure of treatments*) and another referring to the classification of experimental units according to their characteristics (*topographic structure*) [1]. He stresses the importance of correctly formulating the structure of the experiment, emphasizing that it must determine the statistical procedures for inferences, in particular the valid estimates of experimental error. This concept of experimental design was taken up only after three decades [2, 3] and has since been considered from a variety of approaches [4, 5, 6, 7]. However, fundamental concepts, such as experimental material, experimental factor, unit factor, experimental unit and experimental error, are not defined or are treated superficially. Moreover, these approaches do not make distinction between treatment factor and intrinsic factor [8]. Consequently, they ignore that intrinsic experimental factors, such as race, site and year, are "partners" of unit factors, composed of extraneous characteristics that constitute relevant classifications of the observation units. Silva proposes a conceptual and methodological basis for the experimental research, rational and coherent with the logical sequence of the experimental research process [9, 10, 11, 12]. Silva makes a review concerning the conceptual and methodological bases of the experiment [13, 14]. The purpose of this article is to present a synthesis of these contributions, focusing on the procedure for generating the experimental design.

2 Conceptual and Methodological Bases

The experiment is a scientific research method for inferences about causal relationships of characteristics of the units of a target population, that is, relationships of a subset of the class of characteristics that express the performance of

* Corresponding author: João Gilberto Corrêa da Silva
Department of Mathematics and Statistics, Federal University of Pelotas, Pelotas, RS, Brazil. Phone: 55-53-33214410.
E-mail: jgcs1804@yahoo.com.br

these units (response characteristics) with the class of characteristics that supposedly affect them (explanatory characteristics) in the presence of the third class of characteristics of these units (extraneous characteristics). The definition of these three classes of characteristics is established by the problem and the scientific hypothesis that originate the experiment and determine its objectives. The achievement of these objectives depends on the correctness and coherence of the sequential process of the experiment, namely: objective of the experiment $\rightarrow$ planning of the experiment $\rightarrow$ structure of the experiment $\rightarrow$ statistical model $\rightarrow$ statistical inference procedures.

The characteristics of these three classes present in the sample constitute the experimental material. The fraction of the experimental material where an independent observation of a response characteristic is performed is the unit of observation of this characteristic. The values of a response characteristic measured in the observation units have two origins: explanatory characteristics and extraneous characteristics. The effects of the extraneous characteristics on the response characteristics constitute the experimental error. The effects of explanatory characteristics are confounded with experimental error. The resource to reduce this confounding and make it unbiased is experimental control, which is performed by control of experimental techniques, local control, statistical control and randomization (Silva, 2020).

The planning of the experiment defines these three classes of characteristics and the relationships between the explanatory characteristics, between the extraneous characteristics and between these two classes of characteristics, which constitute, respectively, the condition structure, unit structure and experiment structure. The specifications of these structures establish the design of the experiment. The condition structure must be established in accordance with the objectives of the experiment, while the unit structure is elaborated according to the availability of experimental material.

2.1 Planning of the condition structure

The planning of the condition structure comprises the choice of the experimental factors, the levels of these factors and the combinations of these levels. The condition structure may comprise one or more factors. In the first case, the structure is called unifactorial; in the second, factorial. The relationship of two factors is crossed if levels of each factor combine with more than one level of the other factor; it is hierarchical or nested if the different levels of one of the factors combine with different subsets of the levels of the other factor; in this case, the first factor is called nest factor and the second, nested factor. A factor structure can include only crossed factor relations, only nested relations, or both crossed and nested relations. In these three alternative situations, the structure is called crossed, hierarchical or mixed, respectively. The association of the levels of an experimental factor with the sample units can be controlled by experimental control, particularly randomization, or be inherent to these units or associated to a relevant extraneous characteristic. In the first case, the factor is called treatment factor; in the second, intrinsic factor.

2.2 Planning of the unit structure

The planning of the unit structure comprises the choice of the observation unit and the definitions of the local control and of the association between the levels of the experimental factors and the observation units. This planning determines classifications of the observation units into themselves, in groups determined by local control and in experimental units of experimental factors. To each of these classifications corresponds a unit factor. The classes of each of these classifications are the levels of the corresponding unit factor. The effect of a unit factor is a component of the experimental error. Unit factor relationships and unit structures are of the same forms as experimental factor relationships and condition structures.

2.3 Planning the experiment structure

The condition structure and the unit structure are associated by the randomization of the levels of treatment factors to units of the experimental material and the presence of the levels of intrinsic factors in these units. The relationship of these two structures constitutes the experiment structure and the experimental design. Thus, in the experimental design there is a correspondence between the levels of experimental factors and the levels of unit factors. The levels of a unit factor associated with an experimental factor are the experimental units of this experimental factor. The number of experimental units with a level of an experimental factor is the number of repetitions of this level. The relationship between an experimental factor and a unit factor can comprise more than one experimental unit for its levels or a single experimental unit for each of its levels. In the second situation the factors are called equivalents or partners. It occurs when the experimental factor is an intrinsic factor or a treatment factor with one repetition of each of its levels. Effects of equivalent factors are completely confounded. This property is highly relevant and should be considered when planning the experiment.

Unit factors stratify the set of extraneous characteristics of the experimental material. As a consequence, the experimental error is decomposed into as many strata as there are unit factors. The fraction of experimental error that corresponds to a unit factor constitutes a stratum of the experimental error, which originate a stratum of the experiment. The experimental error that affects an effect of experimental factor is a fraction of the experimental error made up of a subset of its strata.

It is convenient that the planning of the condition structure and the unit structure be carried out separately. This procedure is recommended so that the experiment structure is correctly expressed, particularly in complex experiments. However, the condition structure is conditioned to the availability of experimental material and the unit structure must be appropriate for the condition structure. Therefore, these two structures are highly interdependent. A rational strategy for generating the experimental design comprises the following sequence of steps:

- Elaborate the condition structure taking into account the restrictions of experimental material;
- Consider alternative unit structures for this condition structure;
- Choose, among these unit structures, the one that, associated with the condition structure, allows more efficient inferences about the effects of the experimental factors relevant to the objectives of the experiment;
- If a satisfactory unit structure is not found, reconsider the sequence of steps 1, 2 and 3.

Steps 1 and 2 can lead to the formulation of different experiment structures. As a general rule, the researcher should choose the experimental design that provides the most information relevant to the objectives of the experiment at the least cost. For this, the following properties of the experiment plan must be taken appropriately into account: repetition, local control, randomization, orthogonality, balance, confounding and efficiency [12]. Desirable properties of inferences about treatment factors also demand that the plan ensures that the experimental design is consistent with the objectives of the experiment and the following requirements are met: estimation of the components of the experimental error that affect the relevant effects of treatment factors, precision - sensitivity to detect important differences of treatment effects, validity - unbiasedness of inferences, simplicity, economy of resources, feasibility, manifestation of the real effects of the treatments, and provision of statistical inference procedures and measures to assess the degree of uncertainty of these inferences [8, 12].

3 Illustration

Consider an experiment on the effect of energy diet on the body development of lambs between weaning and slaughter (at 70 and 152 days, respectively), with diet in the interval [2.4; 3.2] Mcal/kg DM of metabolizable energy in the target population, and the three levels 2.4, 2.8 and 3.2 in the sample. The experiment is carried out with 15 male and 15 female animals of a race, aged close to 70 days. These animals are housed in a facility with 15 pairs of pens for one animal, each pair equipped with common feeder and drinker. As the characteristics of the environment are heterogeneous, the 15 pairs of pens are classified into five homogeneous blocks of three pairs of pens. Then, in each block, the three pairs of pens are assigned at random to the three diets and the two pens of each pair are randomized to two animals, one of each sex. Homogeneous management and measurement techniques are adopted. The response variables live weight, warm carcass weight and carcass yield are measured in each animal (pen) at slaughter.

The described experimental procedure characterizes two experimental factors: diet and sex, and that diet is a treatment factor, pair of pen is its experimental unit, and pen is the unit of observation of the response variables. The experimental factor sex can be a treatment or an intrinsic factor, depending on the extraneous characteristics of the animals. If these extraneous characteristics are irrelevant, implying that differences between animals of the two sexes are essentially due to sex, sex can be considered a treatment factor. This can occur if animals of both sexes come from the same herd and proper experimental control is used, so that the extraneous characteristics of the animals can be assumed to behave as randomized. However, if animals comprise relevant extraneous characteristics that affect considerably the differences between males and female, sex should be considered an intrinsic factor. This occurs, for example, if males and females come from separate herds, implying relevant differences in extraneous characteristics of the environment and management prior to the experimental period. These two alternative situations are considered below.

3.1 Sex is a treatment factor

The set of experimental conditions contains the six combinations of the three levels of the factor diet (D) with the two levels of the factor sex (S), both fixed treatment factors. The factors D and S have a cross relationship. So, the set of experimental factors is {D, S, D^S}. The condition structure is represented by the symbol D∗S and diagram in Figure1a. It can be verified that it is orthogonal.

The set of observation units contains 30 animals, which are the levels of the unit factor pen (U).The pens constitute pairs that are the levels of the unit factor pair of pens (P), and the pair of pens are classified into blocks to form the unit factor block (B). These three factors U, P and B have a nested relationship; therefore, the set of unit factors is {U, P, B}. These factors are uniform and constitute a nested structure, symbolized by B/P/U and represented by diagram in Figure1b. Therefore, the unit structure is orthogonal.

The association between the unit structure and the condition structure establishes a correspondence of the unit factors P and U with the treatment factors D and S, respectively, and constitute the strata of the experiment B, P⪰D and U⪰S (Figure 1c), where P⪰D symbolizes that factor P is finer than factor D. The generated experiment structure comprises the factors B, D and S that have crossed relations and generate the nested factors P≡B^D, D^S, B^S and B^D^S≡U. The effects originated by the last two factors are interactions between experimental factor and unit factor, which are assumed non-existent, and experimental error. For this reason, they can be aggregated into the factor B^D^S. This structure of experiment is symbolized by B∗D∗S and represented by diagram in Figure 1d. It can be shown that it is orthogonal.

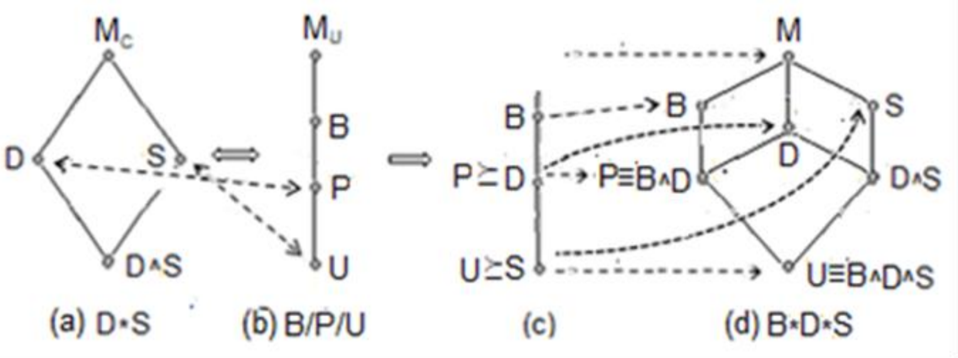

Figure 1 Diagram of the generation of the experiment structure B∗D∗S: (a) condition structure, (b) unit structure, (c) strata of the experiment and (d) experiment structure

The double-headed arrow with a dashed line in Figure 1a, b identifies P and U as the unit factors whose levels are the experimental units of the treatment factors D and S, respectively. It also indicates that the effects of these treatment factors are located in the strata P⪰D and U⪰S, respectively. All information for inferences about the effects of the treatment factor diet comes from the stratum P⪰D, and for inferences about the effects of the treatment factors sex and diet^sex from the stratum U⪰S. In the stratum P⪰D, to different diets correspond different experimental units, but to the same diet correspond ten experimental units. This implies that the variation between experimental units of the observed values of the response variable due to effects of diets is completely confounded with the experimental error within blocks, and this experimental error is partially confounded with diet effects. The fraction of this experimental error not confounded with effects of diets is the variation between pair of pens, excluding variation between blocks and diets. This variation also expresses the interaction between diet and block. However, the absence of interaction between experimental factor and unit factor is a general assumption of the structure of experiment. With this assumption, randomization and the supposition of appropriate experimental control, the variation between pairs of pens, excluding variation between blocks and diets, provides a valid estimate of the variance of the experimental error that affects the effects of diets.

In the stratum U⪰S, to different levels of the factors sex and diet^sex correspond different experimental units (pens), but for each of these factors for a same level correspond fifteen and five experimental units, respectively. This implies that the variation between experimental units due to effects of sex and diet^sex is completely confounded with experimental error within pairs of pens, but this experimental error is partially confounded with effects of these treatment factors. The fraction of this experimental error not confounded with effects of sex and diet^sex is the variation between pens, excluding variation between pair of pens and levels of sex and diet^sex. This variation expresses also the interactions of sex and diet^sex with block, which are assumed non-existent. Because of randomization and appropriate experimental control, this variation provides a valid estimate of the variance of the experimental error that affects the effects of sex and diet^sex.

3.2 Sex is an intrinsic factor

The experimental factor diet is still a treatment factor. Thus, the considerations regarding the factor diet are the same described previously. The changes originate from the differences due to the experimental factor sex, now an intrinsic factor.

The set of observation units comprises 30 animals, as previously. Also, the set of unit factors includes the factors U: pen, P: pair of pens and B: block; but now has an additional factor composed by the extraneous characteristics of the animal whose effects are completely confounded with effects of the intrinsic factor sex. This unit factor, which has two levels, is called herd and denoted by H. Each level of this factor is the set of the extraneous characteristics of the animals of a sex. Factor P is nested in factor B and both are crossed with H, and these three factors nest factor U. Thus, the set of unit factors is {U, P, B, H, P^H, B^H}, where H is equivalent to S. The unit structure is represented by the symbol ((B/P)∗H)/U and represented by diagram in Figure 2b. It can be shown that it is orthogonal.

The equivalence of the experimental factor S and the unit factor H identifies the herd as the experimental unit of the intrinsic factor sex. Therefore, there is only one experimental unit for each level of the intrinsic factor sex.

The structure of the experiment is defined by the association of the unit factors P and H with the experimental factors D and S, respectively, which generate the strata P⪰D and H≡S (Figure 2c). Factors B, D and H≡S are crossed, P is nested in B and D, D^ (H≡S) is nested in D and H≡S, and B^D^ (H≡S) is nested in these factors. It comprises the factors B, D, P≡B^D, H≡S, D^ (H≡S), B^ (H≡S) and B^D^ (H≡S), where the last two factors, as they are combinations of unit factor and experimental factor, can be aggregated in factor B^D^ (H≡S). The experiment structure is represented by the symbol B∗D∗(H≡S) and diagram in Figure 2d. It can be shown to satisfy the requirements for orthogonality.

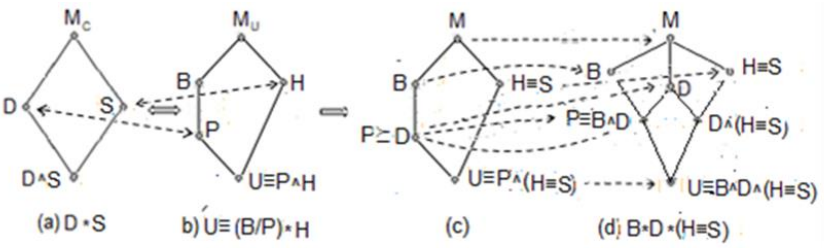

Figure 2 Diagram of the generation of the experiment structure B∗D∗(H≡S): (a) condition structure, (b) unit structure, (c) strata of the experiment and (d) experiment structure

Figure 2a,b identifies P and H as the unit factors whose levels are the experimental units of the experimental factors D and S, respectively, and P⪰D and H≡S as the strata of the experiment where the effects of these experimental factors are located. The considerations regarding the estimation of the variance of the experimental error that affects the effects of the treatment factor diet are the same as in situation 1. In the stratum H≡S the effects of the experimental factor S are completely confounded with effects of the unit factor H. Therefore, this experiment structure does not provide a valid estimate of the variance of the experimental error that affects effects of the experimental factor sex. Therefore, inferences about effects of sex cannot be derived. For a similar reason, inferences about the interaction of sex with diet cannot be derived either.

4 Conclusion

- The construction of the experiment design to achieve the objectives of the experiment requires:

- The clear discrimination between experimental factors and unit factors, and the identification of intrinsic experimental factors.
- The separate formulations of the condition structure and the unit structure, and the association of these two structures consistent with the objectives of the experiment.
- The use of experimental control to achieve the appropriate precision and validity of the inferences.

* These requirements lead to the identification of the strata of the experiment where the experimental factors are located and the experimental units of these factors. It also allows identifying the confounding of the effects of experimental factors and unit factors, and of the components of the experimental error that affect the effects of relevant experimental factors.
* This information makes it possible to outline the path to valid inferences about effects of treatment factors and clarifies the impossibility of valid inferences about effects of intrinsic factors.

This approach allows the generation of an efficient experimental design both from a theoretical and a practical point of view.

Compliance with ethical standards

Acknowledgments

The author thanks the students and researchers who, in his teaching and consulting activities, encouraged and participated in the development of ideas about scientific research, particularly experimental research, presented in this article.

Disclosure of conflict of interest

If two or more authors have contributed in the manuscript; the conflict of interest statement must be inserted here.

References

[1] Fisher RA. Contribution to a discussion of F.Yates' paper on Complex Experiments. London: Journal of the Royal Statistical Society, Supplement. 1935; 2: 229-231.

[2] Nelder JA. The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. London: Proceedings of the Royal Society, Series A. 1965; 273: 147-162.

[3] Nelder JA. The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. London: Proceedings of the Royal Society, Series A. 1965; 273: 163-178.

[4] White RF. Randomization and the analysis of variance. Biometrics. 1975; 31:552-572.

[5] Houtman AM. The analysis of designed experiments [Ph. D. Thesis]. Princeton: Princeton University; 1980.

[6] Brien CJ. Analysis of variance tables based on experimental structure. Biometrics. 1983; 39: 51-59.

[7] Bailey RA. Design of comparative experiments. Cambridge: Cambridge University Press; 2008.

[8] Cox DR. Planning of Experiments. New York: John Wiley; 1958.

[9] Silva JGC. A consideração da estrutura das unidades em inferências derivadas do experimento. Brasilia: Pesquisa Agropecuária Brasileira. 1999; 34: 911-925.

[10] Silva JGC. A estrutura do experimento e o modelo estatístico [Internet]. Lisbon: Sociedade Portuguesa de Estatística; 2004. Accessed 06 March 2022. Available: https://www.academia.edu/1205906/The_structure_of_the_experiment_and_the_statistical_model.

[11] Silva JGC. A conceptual basis and a new approach to the planning of experiments. Cambridge: Isaac Newton Institute for Mathematical Sciences, University of Cambridge; 2008. Accessed 06 March 2022. Available: https://independent.academia.edu/JgdaSilva/Conferences,-Talks,-Workshops.

[12] Silva JGC Estatística Experimental: Planejamento de Experimentos. 2nd ed. Pelotas, Brazil: Instituto de Física e Matemática, Universidade Federal de Pelotas; 2007. Accessed 06 March 2022. Available: https://www.researchgate.net/publication/307890138_Estatistica_Experimental_Planejamento_de_Experimentos_Experimental_Statistics_Planning_of_Experiments.

[13] Silva JGC. Experiment: Conceptual basis. Delhi/London: Journal of Experimental Agriculture International; 2020; 42(6):7-22.

[14] Silva JGC. Experiment: Methodological basis. Raipur, India: International Journal of Science and Research. 2022; 11(3):1085-1097.