

A Framework for the Statistical Analysis of Fixed Effects Models Using *R*, Illustrated with an Example in Biometry

Teresa A. Oliveira[1,2] and Maria Filomena Teodoro[3,4(✉)]

[1] CEAUL, Faculdade de Ciencias, Universidade de Lisboa, 1649-014 Lisbon, Portugal
[2] Universidade Aberta, Departamento de Ciências e Tecnologia, LE@D – Laboratório de Educação a Distância e eLearning, Lisboa, Portugal
[3] CINAV, Center of Naval Research, Naval Academy, Portuguese Navy, 2810-001 Almada, Portugal
mteodoro64@gmail.com
[4] CEMAT, Center for Computational and Stochastic Mathematics, Instituto Superior Tecnico, Lisbon University, 1048-001 Lisbon, Portugal

Abstract. To ensure meaningful insights experiments must be thoughtfully designed to support structured comparisons across treatments. A key tool in this context is the one-way ANOVA, which evaluates whether observed differences in group means reflect true treatment effects rather than random variation. Exploring the importance of One-Way ANOVA in biometrical studies, discussing its assumptions, applications and advantages, a structured and accessible framework is presented for the statistical analysis of fixed effects models using R, illustrated with an application in biometry.

Designed to support both learners and practitioners, the framework combines clear pedagogical guidance with methodological rigor. It introduces key concepts such as model specification, estimation, and inference, while emphasizing reproducibility and interpretability through annotated ***R*** code considering a practical example. Beyond its educational value, the framework contributes a practical and computationally efficient approach to hypothesis testing, model validation, and post-hoc analysis. By bridging statistical theory with applied research, a valuable resource is here provided for students, educators, and biometricians seeking to conduct robust and transparent fixed effects analyses.

1 Introduction

The ability to draw valid and reproducible conclusions depends fundamentally on the quality of the experimental design [1–4]. When comparing multiple treatments or conditions, rather than performing multiple pairwise comparisons, which would be inefficient and statistically unsound, the One-Way ANOVA offers a more rigorous and reliable solution. Structured experimental frameworks are essential to ensure that observed differences are attributable to the factors under study, rather than to random variation. Beyond one-way analysis of variance (ANOVA), there is a fixed effects model that,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16760, pp. 132–146, 2027.
https://doi.org/10.1007/978-3-032-30521-3_9

under controlled conditions, allow the researchers to assess whether the means of different groups differ significantly [1, 5]. The one-way layout ANOVA is widely applied in biological research once it is particularly suited for experiments where a single categorical factor (e.g., treatment type, genotype, or environmental condition) is used to classify observations into distinct groups. Under the assumption of fixed treatment effects, this model partitions the total variability in the data into components attributable to between-group and within-group variation, enabling formal hypothesis testing regarding group differences [6, 7].

Despite its conceptual simplicity, the correct application and interpretation of fixed effects models require careful attention to model assumptions, estimation procedures, and post-hoc analysis. A structured and reproducible framework for the statistical analysis of fixed effects models using the R programming language, including an illustrative example, is presented. The goal is to provide both a pedagogical resource for learners and a practical toolkit for researchers conducting experimental studies in the life sciences [8–14].

2 Methodological Foundations of ANOVA One-Way Layout: The Fixed Effects Model

The One-Way ANOVA methodology [1, 6–8], originally developed by Sir Ronald A. Fisher in the early 20th century, revolutionized experimental design and statistical inference. The resulting F-statistic follows the F-distribution, named in Fisher's honor, and serves as the basis for drawing conclusions about the presence or absence of significant differences among group means.

One-Way ANOVA is based on several key assumptions: normality, homoscedasticity (equal variances), and independence of observations. These assumptions ensure the validity of the F-test, which is used to determine whether the observed differences among group means are statistically significant. The method involves partitioning the total variability in the data into two components: between-group variability and within-group variability. The F-ratio, calculated as the ratio of between-group variance to within-group variance, serves as the test statistic. A significantly large F-value suggests that the group means are not all equal.

In experiments with one factor the data or observations consist of k independent groups of samples (levels) and in each group n_i measurements are taken.

The response observed for each treatment (or observations in the i-th sample) is a random variable Y_{ij} whose concrete values are represented by y_{ij}, with $i = 1,2,\ldots,k$; $j = 1,2,\ldots, n_i$.

If the number of observations n_i is the same, the samples have all the same dimension and that is known as the balanced case. In general model we consider the existence of samples with different dimensions, this the case designated as unbalanced.

2.1 The One-Way Fixed Effects Model and the Data Structure

The appropriate model to describe this type of data is a linear statistical model of the type:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij},\ \mathrm{i} = 1, 2, \ldots, \mathrm{k};\ \mathrm{j} = 1, 2, \ldots n_i,$$

where:

$Y_{ij} \rightarrow$ random variable representative of j-th observation of i-th treatment;

$\mu \rightarrow$ Parameter common to all treatments and designated by global mean;

$\tau_i \rightarrow$ Single parameter for treatment i designated by the effect of i-th treatment;

$\varepsilon_{ij} \rightarrow$ Random component of the error or residuals.

The main objective of the problem is to compare the mean values $\tau_1, \tau_2, ..., \tau_k$ in order to investigate the existence of significant differences between them, and the analysis can be extended to the estimation of the effects of the treatments.

- In order to carry out the hypothesis tests, the following assumptions must be made regarding the random component:
- The errors, ε_{ij}, must be independent normal random variables, of null mean value and variance σ^2, constant for all factor levels.

The one-way ANOVA can be adapted for the investigation of two different situations regarding the effects of the treatments:

- In the first case the experimenter can initially define his interest in testing only the specific treatments in question and, obviously, draw conclusions only with respect to the treatments under study. These conclusions cannot subsequently be adapted or extrapolated to other similar treatments. The investigator can go further and obtain estimators for the parameters of the model, namely μ, τ_i e σ^2. This type of situation is known as the Fixed Effects Model.
- In a second case, one can consider that k treatments are obtained from a wide population of treatments. These situations commonly occur in more complex experiments, involving a large number of treatments to be compared, the study of the entire population involved being neither practical nor economically feasible. In this case it will be desirable to extend the conclusions drawn from the sample to all treatments of the population, whether involved in the analysis or not. The random variables are represented by τ_i and the objective is more related with testing the variability. This type of situation is known by Random Effects Model.

This paper is devoted to the framework of the first case.

The data structure typically consists of observations of a problem with o factor can be arranged in a table with k columns and n_i rows.

Typical data structure for experiments with one factor and k levels

Treatments					
Observations	**1**	**2**	**3**	**…**	**k**
	y_{11}	y_{12}	y_{13}	⋯	y_{1k}
	y_{21}	y_{22}	y_{23}	⋯	y_{2k}
	⋮	⋮	⋮	⋮	⋮
	$y_{n_1 1}$	$y_{n_2 2}$	$y_{n_3 3}$	⋯	$y_{n_k k}$

2.2 The Fundamentals of Statistical Analysis

In the fixed effects model the effects of treatments τ_i are usually defined as deviations from the global mean and it is assumed that:

$$\sum_{i=1}^{k} \tau_i = 0$$

Consider that:
y_{io} is the sum of the total of observations of ith treatment;
$y_{i\bullet}$ is the mean of observations of i-th treatment;
$y_{\bullet\bullet}$ is the global mean of observations.
Thus:

$$y_{io} = \sum_{j=1}^{n_i} y_{ij}, \, y_{i\bullet} = \frac{y_{io}}{n_i}, \; i = 1, 2, ..., k$$

$$y_{oo} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}, \, y_{\bullet\bullet} = \frac{y_{oo}}{n}, \, n = \sum_{i=1}^{k} n_i$$

The mean, or expected value, for the ith treatment is given by:

$$E(Y_{ij}) = \mu_i = \mu + \tau_i, \, i = 1, ..., k.$$

Consequently, the mean of the ith treatment consists of the sum of the global mean with the effect of the ith treatment. The hypothesis to be tested are:
$H_o : \mu_1 = \mu_2 = ... = \mu_k$, versus $H_1 : \mu_i \neq \mu_j$, for at least one pair (i, j).
Note that, under $H_o : \mu_1 = \mu_2 = ... = \mu_k$, all treatments have μ as common mean.
Another equivalent way of writing these hypotheses is:
$H_o : \tau_1 = \tau_2 = ... = \tau_k = 0$, versus $H_1 : \tau_i \neq 0$, for at least some i.
Note that, testing the equality of treatments, or testing that the effects of treatments are null, is the same. As already mentioned, the appropriate procedure for testing the equality of treatments is called analysis of variance and, the general procedures will be described below.
The designation "Analysis of variance" is due to the fact that all this analysis is based on a partition of the total variability of observations in components. The total sum of squares, represented by SQT, measures the overall variability of the data.
Thus:

$$SQT = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - y_{\bullet\bullet})^2$$

Note that, if SQT is divided by the number of corresponding degrees of freedom, that is, $n_1 + n_2 + ... + n_k - 1 = n - 1$, the obtained result is the estimated variance for the sample taking into account the n observations.

The total sum of squares can be partitioned as follows:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-y_{\bullet\bullet})^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}[\,(y_{i\bullet}-y_{\bullet\bullet})+(y_{ij}-y_{i\bullet})\,]^2$$

($y_{i\bullet}$ is simply added and subtracted from the expression and so the result does not change).

Developing this last expression as square of sums, it comes:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}[\,(y_{i\bullet}-y_{\bullet\bullet})+(y_{ij}-y_{i\bullet})\,]^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{i\bullet}-y_{\bullet\bullet})^2$$
$$+2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{i\bullet}-y_{\bullet\bullet})(y_{ij}-y_{i\bullet})+\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-y_{i\bullet})^2$$

Note that:

$$2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{i\bullet}-y_{\bullet\bullet})(y_{ij}-y_{i\bullet})=0.$$

And, moreover,

$$\sum_{j=1}^{n_i}(y_{ij}-y_{i\bullet})=y_{io}-n_iy_{i\bullet}=y_{io}-n_i\frac{y_{io}}{n_i}=y_{io}-y_{io}=0$$

The total variability can be written as:

$$\sum_{\mathrm{i}=1}^{\mathrm{k}}\sum_{\mathrm{j}=1}^{\mathrm{n}_i}(\mathrm{y}_{\mathrm{ij}}-\mathrm{y}_{\bullet\bullet})^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{i\bullet}-y_{\bullet\bullet})^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-y_{i\bullet})^2$$

Resuming, the total variability can be expressed as a partition of:

(i) The sum of squares of differences between the means of the samples (or treatments) and the global mean;

and

(ii) The sum of squares of differences between observations within the samples (or treatments) and average treatments.

Thus, the difference between the means of treatments observed and the overall mean is a measure of the differences between treatment means.

Symbolically, the above expression can be written as SQT = SQA + SQE, where:

SQT – denotes the total sum of squares

SQA – denotes the sum of squares due to the treatments

SQE – denotes the sum of squares of errors or residuals

The data consist of n observations in total, where SQT has n – 1 degrees of freedom; there are k levels of the factor (k different treatments or samples to compare) and thus

SQA has k − 1 degrees of freedom. As for the error, it is known that for each treatment there is $n_i - 1$ degrees of freedom. Since there are k treatments, the number of degrees of freedom for the error is given by:

$$(n_1 - 1) + (n_2 - 1) + ... + (n_k - 1) = \sum_{i=1}^{k} n_i + \sum_{i=1}^{k} 1 = n - k.$$

The basic identity of the analysis of variance consists of two estimates of variance, one based on the variability within the treatments and another variability between treatments.

If there are no differences between the means of the treatments, these two estimates should be very similar and if they are not, this is a reason to suspect that the observed differences will be caused by differences in the means of the treatments. Consider the following quantities:

$$QMA = \frac{SQA}{k - 1} \quad \text{and} \quad QME = \frac{SQE}{n - k},$$

so that QMA represents the samples mean square and QME represents the error mean square. The hypothesis test for differences between treatment means can be established by comparing the variance estimators QMA and QME. The ratio between QMA and QME is therefore called the variance ratio.

As for the errors, the normality, independence, homoscedasticity and expected null value were assumed, for the variables representative of the observations, Y_{ij}, it is assumed that they are identically distributed, mutually independent, and follow approximately the normal distribution, with mean $\mu + \tau_i$, and have the same variance σ^2 [9–12].

Once that the sum of squares for the total is a sum of squares of normal random variables, then,

$\frac{SQT}{\sigma^2} \approx \chi^2_{n-1}$, $\frac{SQE}{\sigma^2} \approx \chi^2_{n-k}$, $\frac{SQA}{\sigma^2} \approx \chi^2_{k-1}$ (considering H_0 is true).

However, once that SQT = SQA + SQE, then the three sums of squares are not independent. It is necessary to resort to a Theorem, Cochran's Theorem, to establish the independence of SQA and SQE. Thus, if H_0 is true, F_0 follows the F distribution with k – 1, n – k degrees of freedom and is defined by:

$$F_0 = \frac{SQA/k - 1}{SQE/n - k} = \frac{QMA}{QME}$$

F_0 is the statistic to test the hypothesis that there are no differences between the means of the treatments. The decision rule is to reject H0 at the significance level α, if $F_0 > F_{k-1,n-k}(\alpha\%)$. The correspondent ANOVA table to test one factor for fixed effects models is:

ANOVA one-way, fixed effects model: general case

O.V	G.L	S.Q	Q.M	R.V
Between samples (treatments)	k – 1	$SQA = \sum_{i=1}^{k} n_i y_{i\bullet}^2 - n y_{\bullet\bullet}^2$	QMA	$F_0 = \frac{QMA}{QME}$
Residual	n – k	SQE = SQT-SQA	QME	
Total	n – 1	$SQT = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - n y_{\bullet\bullet}^2$		

In computational terms it is simpler to use the expressions presented in the previous table for the calculation of SQT, SQA and SQE, respectively. These expressions are easily deduced from the corresponding equivalent expressions given above. In the balanced case, where the samples have equal size, the data can be arranged analogously in a table with r lines (fixed number of observations per treatment or sample) and k columns (number of treatments or samples) and the computations can be simplified as follows:

Observations in experiments with one factor – balanced case

Treatments					
Observations	**1**	**2**	**3**	**…**	**K**
	y_{11}	y_{12}	y_{13}	…	y_{1k}
	y_{21}	y_{22}	y_{23}	…	y_{2k}
	⋮	⋮	⋮	⋮	⋮
	y_{r1}	y_{r2}	y_{r3}	…	y_{rk}

$$y_{io} = \sum_{j=1}^{r} y_{ij} \text{ and } y_{i\bullet} = \frac{1}{r} \sum_{j=1}^{r} y_{ij}$$

$$SQA = r \sum_{i=1}^{k} y_{i\bullet}^2 - n y_{\bullet\bullet}^2 \text{ with } SQT = \sum_{i=1}^{k} \sum_{j=1}^{r} (y_{ij} - y_{\bullet\bullet})^2, \text{ and } r \; k = n.$$

Thus, the ANOVA table can be simplified:

ANOVA One-way, fixed effects model: balanced case

O.V	G.L	S.Q	Q.M	R.V
Between samples (treatments)	k – 1	$SQA = r \sum_{i=1}^{k} y_{i\bullet}^2 - n y_{\bullet\bullet}^2$	QMA	$F_0 = \frac{QMA}{QME}$
Residual	n – k	SQE = SQT-SQA	QME	
Total	n – 1	SQT		

2.3 Statistical Inference for Error

While QME is used to estimate σ^2, QMA is also an estimator of σ^2 under the null hypothesis, that there are no differences between treatments (H0:$\tau_i = 0$).

Consider the expected values for the mean squares of the error:

$$E(QME) = E\left(\frac{SQE}{n-k}\right) = \frac{1}{n-k}E\left[\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij} - y_{i\bullet})^2\right] = \frac{1}{n-k}E\left[\sum_{i=1}^{k}\sum_{j=1}^{r}y_{ij}^2 - \frac{1}{r}\sum_{i=1}^{k}y_{io}^2\right].$$

Note that $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ and that, $y_{io} = \sum_{j=1}^{r}\mu + \tau_i + \varepsilon_{ij}$, and so:

$$E(QME) = \frac{1}{n-k}E\left[\sum_{i\,=\,1}^{k}\sum_{j=1}^{r}(\mu + \tau_i + \varepsilon_{ij})^2 - \frac{1}{r}\sum_{i=1}^{k}\left(\sum_{j=1}^{r}\mu + \tau_i + \varepsilon_{ij}\right)^2\right]$$

After some lengthy calculations, developing the squares and taking the expected values of the terms in parentheses, and replacing the terms of the square of the error by σ^2, it finally turns out to be:

$$E(QME) = \frac{1}{n-k}\left[n\mu^2 + r\sum_{i=1}^{k}\tau_i^2 + n\sigma^2 - n\mu^2 - r\sum_{i=1}^{k}\tau_i^2 - k\sigma^2\right] = \sigma^2.$$

Similarly it can be proved that $E(QMA) = \sigma^2 + \frac{r\sum_{i=1}^{k}\tau_i^2}{k-1}$.

Finally, in cases for which the analysis of variance test is significant, multiple comparisons can be performed in order to analyze such differences.

2.4 The Scheffé Multiple Comparisons Test

For the analysis of the multiple comparisons, the statistic of Scheffé test is given by:

$T_S = \frac{\left|\hat{\tau}_i - \hat{\tau}_{i'}\right|}{\sqrt{\frac{SQE}{n-k}\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)}}$, which is equivalent to use: $T_S = \frac{\left|\hat{\tau}_i - \hat{\tau}_{i'}\right|}{\sqrt{QME\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)}}$,

where $T_S \approx \sqrt{(k-1)F_{k-1,n-k}(\alpha)}$, and $F_{k-1,n-k}(\alpha)$ represents the value of the distribution F with (k – 1, n – k) degrees of freedom and significance level α.

The null hypothesis can be written as:

$H_0 : \tau_i = \tau_{i'}$,

and each comparison case is rejected if $T_S \geq \sqrt{(k-1)F_{k-1,n-k}(\alpha)}$, or:

$$|\hat{\tau}_i - \hat{\tau}_{i'}| \geq \sqrt{(k-1)F_{k-1,n-k}(\alpha)\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right)\frac{SQE}{n-k}}$$

Thus, the differences between observed treatments higher than the value of the second member of the inequality are significant.

2.5 The Estimation of the Model Parameters

To estimate the general model parameters of:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij},\ i = 1, ..., k\ ,\ j = 1, ..., n_i,$$

the Least Squares Method (LSM) is used.

Considering LSM for the estimation of μ and τ_i, it is not necessary to assume the usual assumptions of normality of the errors. The entity leading to the least squares is denoted by W and is represented by the following expression:

$$W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\varepsilon_{ij}^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \mu - \tau_i)^2$$

The values that minimize W, are obtained by solving the following system of equations:

$$\begin{cases} \dfrac{\partial W}{\partial \mu} = 0 \\ \dfrac{\partial W}{\partial \tau_i} = 0, \quad i = 1, 2, ..., k \end{cases}$$

Thus,

$$\begin{cases} \dfrac{\partial W}{\partial \mu} = 0 \\ \dfrac{\partial W}{\partial \tau_i} = 0, \quad i = 1, 2, ..., k \end{cases} \Leftrightarrow \begin{cases} -2\displaystyle\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0 \\ -2\displaystyle\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0, \quad i = 1, 2, ..., k \end{cases}$$

And, after some simplifications, it can be deduced that:

$$\begin{cases} n\hat{\mu} + n_1\hat{\tau}_1 + n_2\hat{\tau}_2 + ... + n_k\hat{\tau}_k = y_{00} \\ n_1\hat{\mu} + n_1\hat{\tau}_1 \qquad\qquad\qquad\qquad = y_{10} \\ n_1\hat{\mu} \qquad + n_2\hat{\tau}_2 \qquad\qquad\qquad = y_{20} \\ \vdots \\ n_1\hat{\mu} \qquad\qquad\qquad + n_k\hat{\tau}_k = y_{k0} \end{cases}$$

These $k + 1$ equations at $k + 1$ unknowns, are known by least squares normal equations or simply by normal equations. The normal equations are not linearly independent, and the solution for $\mu, \tau_1, ..., \tau_k$ is not unique. The way to overcome this problem is to consider that $\sum_{i=1}^{k}\hat{\tau}_i = 0$, and, based on it, the solution for the normal equations is given by:

$$\hat{\mu} = y_{\bullet\bullet}$$

$\hat{\tau}_i = y_{i\bullet} - y_{\bullet\bullet}$, $i = 1, 2, ..., k$.

It may also be interesting to deduce a confidence interval to estimate the mean of the ith treatment and $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = y_{i\bullet}$, i = 1,2,…,k, is an unbiased estimator for μ_i.

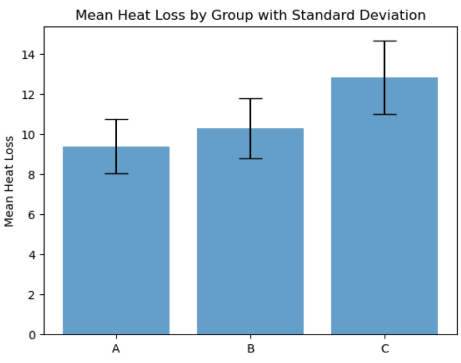
3 Illustrative Example: Classical Way and Using *R*

As an example, consider the case of a pharmaceutical laboratory conducting an experiment with the aim of comparing the heat losses of a particular type of cylinder doors, under three different temperatures and, if significant differences are confirmed, then to clarify which. Thus, these temperatures were tested in some doors, and the following data (in codified form) were obtained:

Temperatures		
A	**B**	**C**
10	12	12
11	12	12
8	11	15
10	8	10
8	10	14
	9	14
	10	

To solve this comparison problem as a first approach it is recommended to have the data visualization, and ***R*** can be used to illustrate the data graphically. Bellow, there is the representation of the ***R*** code and the resultant figure for these data. The visualization of this figure clearly allow the perception that temperature C has a higher average heat loss than the others. Is this perception significant?

```
R Console
> # Dados
> groups <- c("A", "B", "C")
> means <- c(9.40, 10.29, 12.83)
> std_devs <- c(1.34, 1.50, 1.83)
>
> # Criar o gráfico de barras com barras de erro
> barplot(
+   height = means,
+   names.arg = groups,
+   ylim = c(0, max(means + std_devs) + 1),
+   col = "skyblue",
+   main = "Mean Heat Loss by Group with Standard Deviation",
+   ylab = "Mean Heat Loss"
+ )
>
> # Adicionar barras de erro
> arrows(
+   x0 = 1:3,
+   y0 = means - std_devs,
+   x1 = 1:3,
+   y1 = means + std_devs,
+   angle = 90,
+   code = 3,
+   length = 0.1
+ )
```



The main aim is to compare the effect of the three different temperatures, A, B and C, in order to ascertain whether there are significant differences in the heat loss of a particular type of door. Assuming the usual assumptions of normality, independence and homoscedasticity, the proper model for the statistical analysis of these data is the fixed effects model, and the hypothesis, null and alternative, correspond to:

$H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_1 : \mu_i \neq \mu_j$, for at least one pair (i,j).

Temperatures		
A	**B**	**C**
$y_{1o} = 47$	$y_{2o} = 72$	$y_{3o} = 77$
$y_{1\bullet} = \frac{47}{5} = 9.4$	$y_{2\bullet} = \frac{72}{7} = = 10.286$	$y_{3\bullet} = \frac{77}{6} = = 12.833$

$y_{\bullet\bullet} = 10.889$; $n_1 = 5$; $n_2 = 7$; $n_3 = 6$.

$$SQA = \sum_{i=1}^{k} n_i y_{i\bullet}^2 - n y_{\bullet\bullet}^2 = 5(9.4)^2 + 7(10.286)^2 + 6(12.833)^2 - 18(10.889)^2$$

$$= 2170.528 - 18(118.570) = 36.262$$

$$SQT = \sum\sum y_{ij}^2 - n y_{\bullet\bullet}^2 = 449 + 754 + 1005 - 18(10.889) = 73.734.$$

$SQE = SQT - SQA = 73.734 - 36.262 = 37.47$

$SQE = SQT - SQA = 73.734 - 36.262 = 37.472$

ANOVA

Variation Source	Degrees of Freedom	Sum of Squares	Mean squares	Variance ratio
Between temperatures	2	36.262	18.131	7.25
Residual	15	37.472	2.498	
Total	17			

*significance level at 5%, $F_{2,15}(5\%) = 3.68$.

By observing the ANOVA table, and once that the observed value for the variance ratio is higher than the F table value, *7.25 > 3.68*, then the null hypothesis of equality between means is rejected at the *5%* significance level. At this significance level, there is statistical evidence of significant differences between temperatures in terms of heat loss.

To perform the respective ANOVA using R, the code and the results are, respectively:

```
> # Dados brutos
> A <- c(10, 11, 8, 10, 8)
> B <- c(12, 12, 11, 8, 10, 9, 10)
> C <- c(12, 12, 15, 10, 14, 14)
>
> # Criar um data frame com os dados em formato longo
> grupo <- factor(c(rep("A", length(A)), rep("B", length(B)), rep("C", length(C))))
> valor <- c(A, B, C)
> dados <- data.frame(grupo, valor)
>
> # Tabela com os valores por célula
> print("Tabela com os valores por grupo:")
```

```
[1] "Tabela com os valores por grupo:"
> print(dados)
   grupo valor
1      A    10
2      A    11
3      A     8
4      A    10
5      A     8
6      B    12
7      B    12
8      B    11
9      B     8
10     B    10
11     B     9
12     B    10
13     C    12
14     C    12
15     C    15
16     C    10
17     C    14
18     C    14
>
> # ANOVA
> anova_resultado <- aov(valor ~ grupo, data = dados)
> summary(anova_resultado)
            Df Sum Sq Mean Sq F value Pr(>F)
grupo        2  36.32  18.158   7.271 0.0062 **
Residuals   15  37.46   2.497
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
>
```

Note that, using ***R*** there is also the advantage of obtaining the correspondent *p-value,* this case, the *p-value* $= 0.0062 < 0.05$, indicates that, at the significance level of *5%*, there are significant differences between temperatures, confirming the previous classical results and conclusions.

The next step is in what concerns the parameters estimation. Thus, by the classic way, manually computing them from the model, the results are:

$$\hat{\mu} = 10.889$$

$$\hat{\tau}_1 = y_{1\bullet} - y_{\bullet\bullet} = 9.4 - 10.889 = -1.489$$

$$\hat{\tau}_2 = y_{2\bullet} - y_{\bullet\bullet} = 10.286 - 10.889 = -0.603$$

$$\hat{\tau}_3 = y_{3\bullet} - y_{\bullet\bullet} = 12.833 - 10.889 = 1.944$$

And, using ***R***, the code and the respective results are:

```
>
> # Criar os dados
> A <- c(10, 11, 8, 10, 8)
> B <- c(12, 12, 11, 8, 10, 9, 10)
> C <- c(12, 12, 15, 10, 14, 14)
>
> # Criar um vetor com os valores e outro com os grupos
> valores <- c(A, B, C)
> grupo <- factor(c(rep("A", length(A)), rep("B", length(B)), rep("C", length(C))))
>
> # Médias por grupo
> tapply(valores, grupo, mean)
       A        B        C
 9.40000 10.28571 12.83333
>
> # Efeitos estimados (diferença da média de cada grupo em relação à média global)
> media_global <- mean(valores)
> efeitos <- tapply(valores, grupo, mean) - media_global
> efeitos
         A          B          C
-1.4888889 -0.6031746  1.9444444
>
```

And, as it was expected, the previous results are confirmed by using R.

Consider to proceed with multiple comparison test, and taking into account the unbalanced data, at the significance level of *5%* the application of S (Scheffé) multiple comparison test is recommended and the respective results follow, both in classical way and using ***R***.

Comparing temperatures A and B, the results are:

- $|\hat{\tau}_1 - \hat{\tau}_2| = |-1.489 - (-0.603)| = 0.886$

$\sqrt{(3-1)\mathrm{F}_{2,15}(5\%)\left(\frac{1}{5}+\frac{1}{7}\right)(2.489)} = 2.511,$

and the conclusion is that here are no significant differences between temperatures A and B in terms of heat losses;

Comparing temperatures A and C, the results are:

- $|\hat{\tau}_1 - \hat{\tau}_3| = |-1.489 - 1.944| = 3.433$

$\sqrt{(3-1)F_{2,15}(5\%)\left(\frac{1}{5}+\frac{1}{6}\right)(2.489)} = 2.596$, and the conclusion is that there are significant differences between temperatures A and C with respect to heat losses;

Comparing temperatures B and C, the results are:

- $|\hat{\tau}_2 - \hat{\tau}_3| = |-0.603 - 1.944| = 2.547$

$\sqrt{(3-1)\mathrm{F}_{2,15}(5\%)\left(\frac{1}{7}+\frac{1}{6}\right)(2.489)} = 2.386;$

and the conclusion is that there are significant differences between temperatures B and C, with respect to heat losses. Resuming the results:

Contrasts	$\lvert\tau_i - \tau_{i'}\rvert$	Critical value (Scheffé test)	Conclusion
A versus B	0.886	2.511	A does not differ of B
A versus C	3.433	2.596	A differs of C
B versus C	2.547	2.386	B differs of C

The final answer to the Laboratory question is that the temperature C significantly differ from the temperatures A and B, which do not differ from one another, with respect to heat losses of the doors, at a significance level of *5%*.

In order to obtain the parameter estimators from an ANOVA model in ***R***, the group means can be extracted using the model.tables() function. For that, after obtaining the summary results of ANOVA in ***R***, the ***R*** code for Scheffé multiple comparisons must be added:

```
> # Estatísticas básicas
> means <- tapply(df$Temp, df$Group, mean)
> n <- table(df$Group)
> MSE <- summary(anova_model)[[1]]["Residuals", "Mean Sq"]
> df_between <- summary(anova_model)[[1]]["Group", "Df"]
> df_within <- summary(anova_model)[[1]]["Residuals", "Df"]
> alpha <- 0.05
> k <- length(means)
>
> # Comparações de Scheffé
> comparacoes <- combn(names(means), 2, simplify = FALSE)
> for (comp in comparacoes) {
+   g1 <- comp[1]
+   g2 <- comp[2]
+   diff <- abs(means[g1] - means[g2])
+   se <- sqrt(MSE * (1/n[g1] + 1/n[g2]))
+   scheffe_crit <- sqrt((k - 1) * qf(1 - alpha, df_between, df_within)) * se
+   cat(sprintf("%s vs %s:\n", g1, g2))
+   cat(sprintf("  Diferença de médias: %.2f\n", diff))
+   cat(sprintf("  Valor crítico de Scheffé: %.2f\n", scheffe_crit))
+   cat(sprintf("  Significativo? %s\n\n", ifelse(diff > scheffe_crit, "SIM", "NÃO")))
+ }
```

And the respective output is:

```
A vs B:
  Diferença de médias: 0.89
  Valor crítico de Scheffé: 2.51
  Significativo? NÃO

A vs C:
  Diferença de médias: 3.43
  Valor crítico de Scheffé: 2.60
  Significativo? SIM

B vs C:
  Diferença de médias: 2.55
  Valor crítico de Scheffé: 2.39
  Significativo? SIM
```

These results confirm those obtained previously.

An extension of the presented experimental design under the aim of biometry can be found in [13–21] for eventual interess.

Acknowledgements. The authors are thankful to their respective institutions and universities for providing the necessary research support and facilities. The author Teresa A. Oliveira acknowledges the support of LE@D - Laboratory of Distance Education and eLearning, Universidade Aberta, Portugal, funded by National Funds through the FCT, within the scope of the projects UID/4372/2025 e UID/PRR/04372/2025, and of CEAUL Research Unit, UID/00006/2025.

References

1. Oliveira, T.A.: Estatística Aplicada. Ed. Universidade Aberta, nº287 (2004). ISBN: 972-674-444-X1
2. Branda, M.E., Kunneman, M., Meza-Contreras, A.I., et al.: Shared decision-making for patients hospitalized with acute myocardial infarction: a randomized trial. Patient Prefer. Adherence **16**, 1395–1404 (2022)
3. Eldridge-Smith, E.D., Manber, R., Tsai, S., et al.: Stepped care management of insomnia co-occurring with sleep apnea: the air study protocol. Trials **23**(1), 806 (2022)
4. Fatoye, F., Gebrye, T., Mbada, C.E., et al.: Cost effectiveness of virtual reality game compared to clinic based Mckenzie extension therapy for chronic non-specific low back pain. Bull. Geodesique **16**(6), 601–609 (2022)
5. Lipchik, G.L., Nicholson, R.A., Penzien, D.B.: Allocation of patients to conditions in headache clinical trials: randomization, stratification, and treatment matching. Headache **45**(5)(498), 419–428 (2005)
6. Brown, B.W.J.: Designing for cancer clinical trials: selection of prognostic factors. Cancer Treat. Rep. **64**(2–3), 499–502 (1980)
7. Ross, S.M.: Introduction to Probability and Statistics for Engineers and Scientists, 6th edn. Academic Press, San Diego, CA (2020)
8. Morgado, L., Teodoro, F., Perdicoulis, T.: Manual de Métodos Estatísticos em Ciências Biomédicas. Universidade de Trás os Montes e Alto Douro (UTAD, Vila Real, série Didáctica Cências Aplicadas 400 (2010). ISBN 978-972-669-018-2
9. Larsen, R., Marx, M.L.: Statistics, 5th edn. Prentice-Hall Inc. (2012)
10. Tamhane, A., Dunlop, D.: Statistics and Data Analysis. Prentice-Hall Inc., New Jersey (2000)
11. Thadewald, T., Buning, H.: Jarque-Bera test and its competitors for testing normality: a power comparison. Working Paper No.2004/9. Free University Berlin, School of Business & Economics, Berlin (2004). https://hdl.handle.net/10419/49919. Accessed 14 Apr 2025
12. Pocock, S.J.: Allocation of patients to treatment in clinical trials. Biometrics **35**(1), 183–197 (1979)
13. Pocock, S.J.: Clinical Trials, 5th edn. e-book. John Wiley & Sons, New Jersey (2013)
14. Atkinson, A.C., Biswas, A.: Bayesian adaptive biased-coin designs for bayesian adaptive biased-coin designs for clinical trials with normal responses. Biometrics **61**(1), 118–125 (2005)
15. Gail, M.H., Mark, S.D., Carroll, R.J., et al.: On design considerations and randomization based inference for community intervention trials. Stat. Med. **15**(11), 1069–1092 (1996)
16. Simon, R.: Restricted randomization designs in clinical trials. Biometrics **35**(2), 503–612 (1979)
17. Ling, S.X., Hobbs, B.P., Kaizer, A.M., Koopmeiners, J.S.: Calibrated dynamic borrowing using capping priors. J. Biopharm. Stat. **31**(6), 852–867 (2021)
18. Daniells, L., Mozgunov, P., Bedding, A., Jaki, T.: A comparison of bayesian information borrowing methods in basket trials and a novel proposal of modified exchangeability-nonexchangeability method. Stat. Med. **42**(24), 4392–4417 (2023)
19. Bruce, C.L., Iflaifel, M., Montgomery, A., et al.: Choosing and evaluating randomisation methods in clinical trials: a qualitative study. Clin. Trial **25**(1), 199 (2024)
20. Scott, D., Patterson, B.J.: Bioequivalence and Statistics in Clinical Pharmacology. Chapman and Hall/CRC, New York (2017)
21. Zeng, L., Qureshi, R., et al.: A registration of phase 3 crossover trials on clinicaltrials.gov. Trials **21**(613) (2020)