Considering Type II Error in a Two-Level Factorial Design

Mohammed Majrashi, Estefany Ocegueda, Tanis Sarbatananda

Advisor: Dr. Jessica Jaynes, TA's: Jessica Romero, Jose Toledo, Department of Mathematics - California State University, Fullerton

Objective

The objective of this research is to replicate the methodology found in the paper: *Selecting Relevant Effects in Factorial Designs* by Grima et al.(2020). This methodology takes into account Type II error when determining the factorial effects with the most effect on the response. We implement this methodology with an application to determine the percentage of Myrtle essential oil yield based on several factors from the research paper by Ammar et al.(2010).

Introduction

Researchers deciding on a significant effect often have to watch out for making two types of error, known as α and β which coincide with the probability of committing a Type I or Type II error. To determine whether an effect is truly statistically significant, α is often set at 5% while β is often set anywhere between 10%-20% . To get an accurate depiction of what factorial effects are significant, the probability of committing a type II error needs to be considered.

	Null hypothesis is TRUE	**Null hypothesis is FALSE**
Reject null hypothesis	Type I Error (False positive)	Correct outcome (True positive)
Fail to reject null hypothesis	Correct outcome (True negative)	Type II Error (False negative)

Factorial Designs

- Experimental design is a branch of statistics dealing with the design and analysis of experiments and is commonly used in fields such as medicine, education and human behavior.
- Factorial designs are employed to investigate the effect of multiple factors simultaneously on a response.
- A factorial design with k factors each at two levels consists of all possible combinations: 2^k.
- Factorial effects represent the estimates of the factors, and are known as main effects and interaction effects. A main effect is a quantity of an effect from a single factor while interaction effects are measurements of effects from multiple factors occurring simultaneously.

Methodology: Exploring Type I Error

Various thresholds exist for determining factorial effect significance. Two commonly used thresholds based on Type I error are:

Margin of error (ME):

$$ME = t_{d,\alpha/2} \cdot PSE \tag{1}$$

where $d = N/3$ is the degree of freedom and N is the total number of effects. Factorial effects greater than the ME are seen as "significant," but because these estimates are considered simultaneously, some of them may be incorrectly identified.

Simultaneous margin of error (SME):

$$SME = t_{d,\gamma} \cdot PSE \tag{2}$$

where $\gamma = (1 + 0.95^{1/N})/2$. SME can be used to account for simultaneous testing. A factorial effect larger than SME is declared significant.

Both of these thresholds are a function of the Pseudo Standard Error (PSE), which is calculated as follows:

$$PSE = 1.5 \cdot median_{|\hat{\theta}_i| < 2.5 s_0} |\hat{\theta}_i|$$

$$s_0 = 1.5 \cdot median |\hat{\theta}_i|$$

where $\hat{\theta}_i$ represents the i^{th} factorial effect.

Methodology: Exploring Type II Error

While Type I error is commonly used for determining effect significance, it is also important to consider the probability of dismissing an effect when its influence on the response is relevant (Type II error). In addition to ME and SME to determine factorial effect significance, Grima et al. (2020) introduce the critical value for relevance (CVR) threshold as a means for incorporating Type II error. CVR takes into account the circumstances when the effects are located near the critical value or an effect is disregarded due to its value deviating from other effects.

Three steps for calculating CVR:

- Step 1: Determine a change in response known as Minimum Effect Size of Interest (MESI) based on a specified β value.
- Step 2: Estimate the standard deviation of the effects by recalculating PSE only based on non-significant effects.
- Step 3: Determine the critical value for relevance:

$$CVR = t_{v,d,\beta} \cdot PSE \tag{3}$$

where v is the degrees of freedom equal to the number of non-significant effects and $d = MESI/PSE$, such that PSE is based on Step 2.

Application: Myrtle Essential Oil Yield

In Ammar et al. (2010), a 2^4 full factorial design is used to determine the percentage of Myrtle essential oil yield based on four factors: efficacy of granulometry (A), processing time (B), condensation flow (C), and mass ratio plant: water (D). The table below provides the design with 16 runs as well as the addition of four center runs:

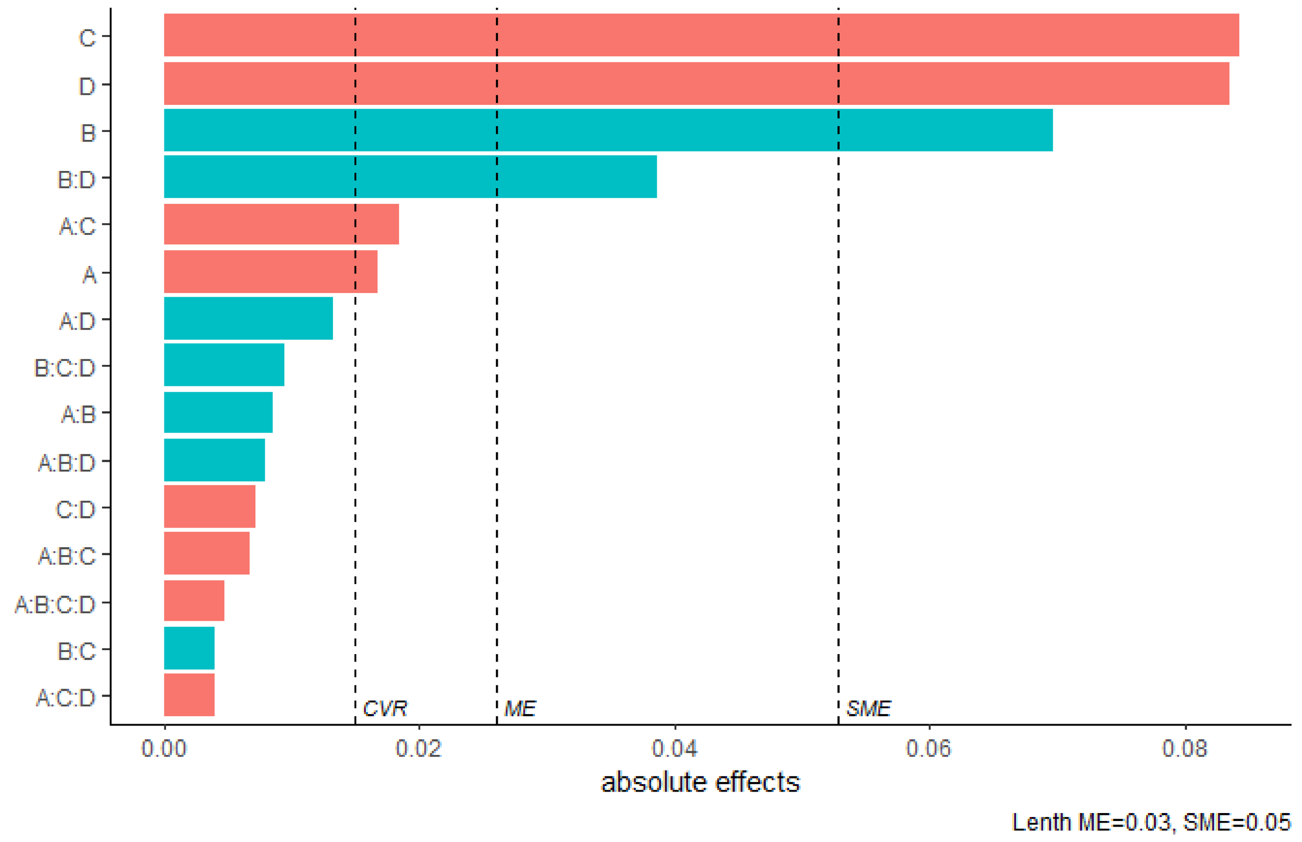
Run	A	B	C	D	Response (%)
1	-1	-1	-1	-1	0.706
2	1	-1	-1	-1	0.688
3	-1	1	-1	-1	0.740
4	1	1	-1	-1	0.727
5	-1	-1	1	-1	0.647
6	1	-1	1	-1	0.604
7	-1	1	1	-1	0.674
8	1	1	1	-1	0.628
9	-1	-1	-1	1	0.596
10	1	-1	-1	1	0.587
11	-1	1	-1	1	0.663
12	1	1	-1	1	0.710
13	-1	-1	1	1	0.502
14	1	-1	1	1	0.471
15	-1	1	1	1	0.619
16	1	1	1	1	0.598
17	0	0	0	0	0.699
18	0	0	0	0	0.644
19	0	0	0	0	0.633
20	0	0	0	0	0.580

For illustration purposes we arbitrary set $\alpha = 5\%$, $\beta = 20\%$, and MESI = 0.025.

Figure 1:Tunisian myrtle (Myrtus communis L.)

Results

The Pareto plot below plots the absolute value of the factorial effects from the Ammar et al. (2010) data, where red bars indicate negative factorial effects and blue bars indicate positive factorial effects.



Discussion

Through the three different thresholds, we find varying factorial effects significant:

- According to the ME threshold, factorial effects C, D, B, and B:D are significant.
- According to the SME threshold, factorial effects C, D, and B are significant
- According to the CVR threshold, factorial effects C, D, B, B:D, A:C, and A are significant

The decision on which threshold to use should be made in conjunction with the subject matter expert. Also, it is important to note that the results will vary depending on the specification of α, β, and MESI.

References

1 Ammar, Ahlem Haj, et al. "Optimization of Operating Conditions of Tunisian Myrtle (Myrtus Communis L.) Essential Oil Extraction by a Hydrodistillation Process Using a 24 Complete Factorial Design." Flavour and Fragrance Journal, vol. 25, no. 6, 2010, pp. 503–507., doi:10.1002/ffj.2011.

2 Grima, Pere, et al. "Selecting Relevant Effects in Factorial Designs." Quality and Reliability Engineering International, vol. 36, no. 7, 2020, pp. 2370–2378., doi:10.1002/qre.2702.

3 Wu, Chien-Fu, and Michael Hamada. Experiments: Planning, Analysis, and Optimization. John Wiley Sons, Inc., 2021.