

ANOVA-nSTAT: ANOVA methodology and computational tools in the paradigm of new statistics

WenJun Zhang¹, YanHong Qi²

¹School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China

²Sun Yat-sen University Library, Sun Yat-sen University, Guangzhou 510275, China

E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 25 July 2023; Accepted 28 August 2023; Published online 17 September 2023; Published 1 March 2024



Abstract

Based on the statistical significance testing, the Analysis of Variance (ANOVA) is one of the most popular statistics widely used in experimental sciences. However, in recent years, the statistical significance testing has been widely criticized for its various shortcomings. To address this problem, in present article we developed the new ANOVA methodology in the paradigm of new statistics. In this methodology, effect size testing is added to six most used ANOVA methods and the finer p -value standardard is used in the statistical significance testing of eight ANOVA methods. Both online and offline computational tools are developed for free use.

Keywords analysis of variance (ANOVA); effect sizes; effect size testing; significance testing; standalone software; online computation.

Computational Ecology and Software
ISSN 2220-721X
URL: <http://www.iaees.org/publications/journals/ces/online-version.asp>
RSS: <http://www.iaees.org/publications/journals/ces/rss.xml>
E-mail: ces@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

The Analysis of Variance (ANOVA) is one of the most popular statistics widely used in experimental sciences including biology, ecology, medicine, and agronomy, etc. The ANOVA theory originated from the British mathematician Fisher, and many ANOVA methods have been developed so far. ANOVA can be used to distinguish various variances according to the causes of variation; use appropriate variances as experimental errors to conduct significance tests, and find best treatment measures, etc. Traditionally, one can use ANOVA to compare whether there are significant differences among treatments and factors, and discover the interactions and causes between treatments or factors (Zhang and Qi, 2005).

ANOVA is based on the statistical significance testing (Fisher, 1935; Zhang, 2022c). Statistical significance tests are one of the most important statistical inference methods in statistics (Fisher, 1935; Yates, 1951; Amrhein et al., 2019; Sellke et al., 2001; Zhang, 2022c). Whether a research result is statistically significant is mainly determined by using the p -value obtained from hypothesis testing (Bergstrom and West,

2021). The p -value is at the heart of the statistical significance testing (Sun, 2016; Zhang, 2022c). In recent years, however, the statistical significance testing have been questioned, mainly because the paradigm of significance tests is wrong, p -value is too sensitive, p -value is a dichotomous subjective index, and statistical significance is related to sample size, etc (Sellke et al., 2001; Trafimow and Marks, 2015; Baker, 2016; Wasserstein and Lazar, 2016; McShane and David, 2017; Amrhein et al., 2019; Tong, 2019; Wasserstein et al., 2019; Zhang, 2022a-c). Statistical significance testing has been one of the sources of false conclusions and research reproducibility crisis (Ioannidis, 2005; Open Science Collaboration, 2015; Errington et al., 2021; Huang, 2021a-b, 2023; Kafdar, 2021; Nature Editorial, 2021; Vrieze, 2021; Zhang, 2022c). In the last year, Zhang (2022c) has reviewed and detailed p -value based statistical significance testing, including concepts, misuses, critiques, and solutions, etc. In recent years the search for better statistics to replace p -value based statistical significance testing is becoming a hot topic in statistics (Huang, 2021a-b, 2023). Many statisticians invoke to abandon p -value based statistical significance testing and replace them with effect size, Bayesian methods, meta-analysis, etc. For example, effect size is considered to be one of the main contents in new statistics (Cumming, 2013; Zhang, 2023). Different from the significance testing, the effect size paradigm guides researchers to pay attention to scientific significance (Cohen, 1988, 2008; Elis, 2010; Lenhard and Lenhard, 2016; Huang, 2021a-b, 2023; Zhang, 2022c). To address these problems in statistics, the construction of a new statistics, including the wide use of effect sizes, is argued. In addition to writing, publishing and adopting new statistical monographs and textbooks, the most urgent task is to revise and distribute various statistical software based on the new statistics for further use (Zhang, 2022a-c, 2023).

In present article, we developed the new ANOVA methodology in the paradigm of new statistics. In this methodology, effect size testing was added to six most used ANOVA methods and the finer p -value standard was used in p -value statistical significance testing of eight ANOVA methods. Both online and offline computational tools are developed for free use.

2 ANOVA (Analysis of Variance): New Statistics

2.1 One-Factor ANOVA

2.1.1 Randomized Complete Design

Assuming that there are no subsamples in the experiment, the data matrix is (x_{ij}) , $i=1, \dots, n; j=1, 2, \dots, m$, where n is the number of treatments and m is the number of replications. Calculate the sum of squares of treatments (SS_t) and the sum of squares of errors (SS_e) as follows

$$SS_t = (\sum_{i=1}^n (\sum_{j=1}^m x_{ij})^2) / m - C$$

$$SS_e = \sum_{i=1}^n \sum_{j=1}^m x_{ij}^2 - SS_t - C$$

where $C = (\sum_{i=1}^n \sum_{j=1}^m x_{ij})^2 / (n * m)$. The total degrees of freedom, the degrees of freedom of treatments, and the degrees of freedom of errors are

$$df = n * m - 1$$

$$df_t = n - 1$$

$$df_e = n * (m - 1)$$

Calculate F -value as the following

$$F = (SS_t / df_t) / (SS_e / df_e)$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c)

$$\eta^2 = Fdf_i / (Fdf_i + df_e)$$

$\eta^2 < 0.10$, no treatment effect; $\eta^2 \in [0.10, 0.30)$, small treatment effect; $\eta^2 \in [0.30, 0.50)$, intermediate treatment effect; $\eta^2 \geq 0.50$, large treatment effect. If there is at least an intermediate treatment effect (or small treatment effect, or large treatment effect, which is dependent upon research's requirement), i.e., $\eta^2 \geq 0.30$, use the effect size measure, Cohen's d , to compare between-mean differences of treatments (Cohen, 1988; Zhang, 2023). For a pair of treatments i and j ($i, j = 1, 2, \dots, n; i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two treatments i and j , respectively, and s_i and s_j are the standard deviation of the two treatments i and j . $|d| < 0.20$, no treatment effect; $|d| \in [0.20, 0.50)$, small treatment effect; $|d| \in [0.50, 0.80)$, intermediate treatment effect; $|d| \geq 0.80$, large treatment effect.

(2) Significance testing

There are significant differences among treatments, if $F > F_\alpha(df_i, df_e)$, $\alpha = 0.001, 0.0001$, etc. If so, use LSD method to compare between-mean differences of treatments, i.e., calculate $S = (2 * (SS_e / df_e) / m)^{1/2}$, thus $LSD_\alpha = t_\alpha(df_e) * S$. Calculate $x_{bi} - x_{bj}$, $i, j = 1, 2, \dots, n; i \neq j$, $x_{bi} = \sum_{j=1}^m x_{ij} / m$. If $(x_{bi} - x_{bj}) \geq LSD_\alpha$, there is a significant difference between treatments i and j .

2.1.2 Randomized Block Design

Assuming that there are no subsamples in the experiment, the data matrix is (x_{ij}) , $i = 1, \dots, n; j = 1, 2, \dots, m$, where n is the number of treatments and m is the number of blocks. Calculate the total sum of squares (SS_T), the sum of squares of blocks (SS_b), the sum of squares of treatments (SS_t), and the sum of squares of errors (SS_e) as follows

$$\begin{aligned} SS_T &= \sum_{i=1}^n (\sum_{j=1}^m x_{ij}^2) - C \\ SS_b &= \sum_{j=1}^m (\sum_{i=1}^n x_{ij}^2) / n - C \\ SS_t &= \sum_{i=1}^n (\sum_{j=1}^m x_{ij}^2) / m - C \\ SS_e &= SS_T - SS_b - SS_t \end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^m x_{ij})^2 / (n * m)$. The degrees of freedom of blocks, the degrees of freedom of treatments, and the degrees of freedom of errors are

$$\begin{aligned} df_b &= m - 1 \\ df_t &= n - 1 \\ df_e &= (n-1) * (m-1) \end{aligned}$$

Calculate F -value as the following

$$F = (SS_t / df_t) / (SS_e / df_e)$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c)

$$\eta^2 = Fdf_i / (Fdf_i + df_e)$$

$\eta^2 < 0.10$, no treatment effect; $\eta^2 \in [0.10, 0.30)$, small treatment effect; $\eta^2 \in [0.30, 0.50)$, intermediate treatment effect; $\eta^2 \geq 0.50$, large treatment effect. If there is at least an intermediate treatment effect (or small treatment effect, or large treatment effect, which is dependent upon research's requirement), i.e., $\eta^2 \geq 0.30$, use the effect size measure, Cohen's d , to compare between-mean differences of treatments (Cohen, 1988; Zhang, 2023). For a pair of treatments i and j ($i, j = 1, 2, \dots, n; i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two treatments i and j , respectively, and s_i and s_j are the standard deviation of the two treatments i and j . $|d| < 0.20$, no treatment effect; $|d| \in [0.20, 0.50)$, small treatment effect; $|d| \in [0.50, 0.80)$, intermediate treatment effect; $|d| \geq 0.80$, large treatment effect.

(2) Significance testing

There are significant differences among treatments, if $F > F_\alpha(df_i, df_e)$, $\alpha = 0.001, 0.0001$, etc. If so, use LSD method to compare between-mean differences of treatments.

2.1.3 Randomized Block Design with Subsamples

Assuming that there are subsamples in the experiment, the data matrix is (x_{ijk}) , $i = 1, \dots, n; j = 1, 2, \dots, h; k = 1, 2, \dots, m$, where n is the number of treatments, m is the number of blocks, h is the number of subsamples. Calculate the total sum of squares (SS_T), the sum of squares of subsamples (SS_b), the sum of squares of blocks (SS_h), the sum of squares of treatments (SS_t), the sum of squares of errors (SS_e), and the sum of squares of sampling errors (SS_s) as follows

$$\begin{aligned} SS_T &= \sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk}^2 - C \\ SS_u &= \sum_{i=1}^n \sum_{k=1}^m (\sum_{j=1}^h x_{ijk})^2 / h - C \\ SS_b &= \sum_{k=1}^m (\sum_{i=1}^n \sum_{j=1}^h x_{ijk})^2 / (n * h) - C \\ SS_t &= \sum_{i=1}^n (\sum_{k=1}^m \sum_{j=1}^h x_{ijk})^2 / (m * h) - C \\ SS_e &= SS_u - SS_b - SS_t \\ SS_s &= SS_T - SS_u \end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk})^2 / (n * h * m)$. The total degrees of freedom, the degrees of freedom of subsamples, the degrees of freedom of blocks, the degrees of freedom of treatments, the degrees of freedom of errors, and the degrees of freedom of sampling errors are

$$\begin{aligned} df_i &= n * h * m - 1 \\ df_u &= n * m - 1 \\ df_b &= m - 1 \\ df_t &= n - 1 \\ df_e &= (n - 1) * (m - 1) \\ df_s &= n * m * (h - 1) \end{aligned}$$

Calculate F -value as the following

$$F=(SS_t/df_t)/(SS_e/df_e)$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c)

$$\eta^2 = Fdf_t / (Fdf_t + df_e)$$

$\eta^2 < 0.10$, no treatment effect; $\eta^2 \in [0.10, 0.30)$, small treatment effect; $\eta^2 \in [0.30, 0.50)$, intermediate treatment effect; $\eta^2 \geq 0.50$, large treatment effect. If there is at least an intermediate treatment effect (or small treatment effect, or large treatment effect, which is dependent upon research's requirement), i.e., $\eta^2 \geq 0.30$, use the effect size measure, Cohen's d , to compare between-mean differences of treatments (Cohen, 1988; Zhang, 2023). For a pair of treatments i and j ($i, j = 1, 2, \dots, n; i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two treatments i and j , respectively, and s_i and s_j are the standard deviation of the two treatments i and j . $|d| < 0.20$, no treatment effect; $|d| \in [0.20, 0.50)$, small treatment effect; $|d| \in [0.50, 0.80)$, intermediate treatment effect; $|d| \geq 0.80$, large treatment effect.

(2) Significance testing

There are significant differences among treatments, if $F > F_\alpha(df_t, df_e)$, $\alpha = 0.001, 0.0001$, etc. If so, use LSD method to compare between-mean differences of treatments.

2.1.4 Nested Design

Assuming that the data matrix is (x_{ijk}) , $i = 1, \dots, n; j = 1, 2, \dots, h; k = 1, 2, \dots, m$, where n is the number of treatments, m is the number of subsamples, h is the number of blocks. Calculate the total sum of squares (SS_T), the sum of squares of treatments (SS_t), the sum of squares of blocks (SS_b), and the sum of squares of errors (SS_e) as follows

$$\begin{aligned} SS_T &= \sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk}^2 - C \\ SS_t &= \sum_{i=1}^n (\sum_{k=1}^m \sum_{j=1}^h x_{ijk})^2 / (m * h) - C \\ SS_b &= \sum_{j=1}^h (\sum_{i=1}^n \sum_{k=1}^m x_{ijk})^2 / m - SS_t - C \\ SS_e &= SS_T - SS_t - SS_b \end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk}^2) / (n * h * m)$. The total degrees of freedom, the degrees of freedom of treatments, the degrees of freedom of blocks, and the degrees of freedom of errors are

$$\begin{aligned} df_t &= n * h * m - 1 \\ df_i &= n - 1 \\ df_b &= n * (h - 1) \\ df_e &= n * h * (m - 1) \end{aligned}$$

Calculate F -value as the following

$$F=(SS_t/df_t)/(SS_e/df_e)$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c)

$$\eta^2 = Fdf_i / (Fdf_i + df_e)$$

$\eta^2 < 0.10$, no treatment effect; $\eta^2 \in [0.10, 0.30)$, small treatment effect; $\eta^2 \in [0.30, 0.50)$, intermediate treatment effect; $\eta^2 \geq 0.50$, large treatment effect. If there is at least an intermediate treatment effect (or small treatment effect, or large treatment effect, which is dependent upon research's requirement), i.e., $\eta^2 \geq 0.30$, use the effect size measure, Cohen's d , to compare between-mean differences of treatments (Cohen, 1988; Zhang, 2023). For a pair of treatments i and j ($i, j = 1, 2, \dots, n; i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two treatments i and j , respectively, and s_i and s_j are the standard deviation of the two treatments i and j . $|d| < 0.20$, no treatment effect; $|d| \in [0.20, 0.50)$, small treatment effect; $|d| \in [0.50, 0.80)$, intermediate treatment effect; $|d| \geq 0.80$, large treatment effect.

(2) Significance testing

There are significant differences among treatments, if $F > F_\alpha(df_i, df_e)$, $\alpha = 0.001, 0.0001$, etc. If so, use LSD method to compare between-mean differences of treatments.

2.2 Two-Factor ANOVA**2.2.1 Two-Factor Classification**

Assuming that the data matrix is (x_{ijk}) , $i = 1, \dots, n; j = 1, 2, \dots, m; k = 1, 2, \dots, r$, where n is the number of levels of factor A, m is the number of levels of factor B, r is the number of replications. Calculate the total sum of squares (SS_T), the sum of squares of factor A (SS_a), the sum of squares of factor B (SS_b), the sum of squares of treatments' combination (SS_t), the sum of squares of interactions between factor A and factor B (SS_{ab}), and the sum of squares of errors (SS_e) as follows

$$\begin{aligned} SS_T &= \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r x_{ijk}^2 - C \\ SS_a &= \sum_{i=1}^n (\sum_{j=1}^m \sum_{k=1}^r x_{ijk})^2 / (m*r) - C \\ SS_b &= \sum_{j=1}^m (\sum_{i=1}^n \sum_{k=1}^r x_{ijk})^2 / (n*r) - C \\ SS_t &= \sum_{i=1}^n \sum_{j=1}^m (\sum_{k=1}^r x_{ijk})^2 / r - C, \\ SS_{ab} &= SS_t - SS_a - SS_b \\ SS_e &= SS_T - SS_t \end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r x_{ijk})^2 / (n*m*r)$. The total degrees of freedom, degrees of freedom of factor A, the degrees of freedom of factor B, the degrees of freedom of treatments' combination, the degrees of freedom of interactions between factor A and factor B, and the degrees of freedom of errors are

$$\begin{aligned} df_T &= n*m*r - 1 \\ df_a &= n - 1 \\ df_b &= m - 1 \\ df_t &= n*m - 1 \\ df_{ab} &= (n-1)*(m-1) \\ df_e &= n*m*(r-1) \end{aligned}$$

Calculate F -values for factor A, factor B, and the interactions between factor A and factor B as follows

$$F_a = (SS_a/df_a)/(SS_e/df_e)$$

$$F_b = (SS_b/df_b)/(SS_e/df_e)$$

$$F_{ab} = (SS_{ab}/df_{ab})/(SS_e/df_e)$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c), for factor A, factor B, and interactions between factor A and factor B

$$\eta_a^2 = F_a df_a / (F_a df_a + df_e)$$

$$\eta_b^2 = F_b df_b / (F_b df_b + df_e)$$

$$\eta_{ab}^2 = F_{ab} df_{ab} / (F_{ab} df_{ab} + df_e)$$

η_a^2 , η_b^2 , or $\eta_{ab}^2 < 0.10$, no effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \in [0.10, 0.30)$, small effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \in [0.30, 0.50)$, intermediate effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \geq 0.50$, large effect.

If there is at least an intermediate interaction effect (or small interaction effect, or large interaction effect, which is dependent upon research's requirement), i.e., $\eta_{ab}^2 \geq 0.30$, ANOVA should be terminated. Otherwise, continue for next steps.

If there is at least an intermediate factor effect (or small factor effect, or large factor effect, which is dependent upon research's requirement), i.e., $\eta_a^2 \geq 0.30$ (or $\eta_b^2 \geq 0.30$), use the effect size measure, Cohen's d , to compare between-mean differences of the factor A (or B) (Cohen, 1988; Zhang, 2023). For a pair of levels i and j ($i, j = 1, 2, \dots, n$; $i \neq j$) of factor A (or B; $i, j = 1, 2, \dots, m$; $i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two levels i and j , respectively, and s_i and s_j are the standard deviation of the two levels i and j . $|d| < 0.20$, no level effect; $|d| \in [0.20, 0.50)$, small level effect; $|d| \in [0.50, 0.80)$, intermediate level effect; $|d| \geq 0.80$, large level effect.

(2) Significance testing

For significance testing, there are significant differences among treatments (levels) of factor A, if $F > F_a(df_a, df_e)$; there are significant differences among treatments (levels) of factor B, if $F > F_b(df_b, df_e)$; there are significant differences among the interactions between factor A and factor B, if $F > F_{ab}(df_{ab}, df_e)$. $\alpha = 0.001$, 0.0001, etc.

If $F \leq F_a(df_a, df_e)$, use LSD method to compare between-mean differences. Calculate

$$S_a = (2 * (SS_e/df_e) / (m * r))^{1/2}$$

$$S_b = (2 * (SS_e/df_e) / (n * r))^{1/2}$$

And the corresponding LSDs are

$$LSD_a = t_a(df_e) * S_a$$

$$LSD_b = t_b(df_e) * S_b$$

Calculate the difference of pairwise means above and if it is greater than the corresponding LSD_{α} , then the difference of means are significant.

2.2.2 Two-Factor Randomized Complete Block Design

Assuming that the data matrix is (x_{ijk}) , $i=1, \dots, n$; $j=1, 2, \dots, m$; $k=1, 2, \dots, r$, where n is the number of levels of factor A, m is the number of levels of factor B, r is the number of replications. Calculate the total sum of squares (SS_T), sum of squares of factor A (SS_a), the sum of squares of factor B (SS_b), the sum of squares of treatments' combination (SS_t), the sum of squares of interactions between factor A and factor B (SS_{ab}), the sum of squares of blocks (SS_u), and the sum of squares of errors (SS_e) as follows

$$\begin{aligned} SS_T &= \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r x_{ijk}^2 - C \\ SS_a &= \sum_{i=1}^n (\sum_{j=1}^m \sum_{k=1}^r x_{ijk})^2 / (m*r) - C \\ SS_b &= \sum_{j=1}^m (\sum_{i=1}^n \sum_{k=1}^r x_{ijk})^2 / (n*r) - C \\ SS_t &= \sum_{i=1}^n \sum_{j=1}^m (\sum_{k=1}^r x_{ijk})^2 / r - C, \\ SS_{ab} &= SS_t - SS_a - SS_b \\ SS_u &= \sum_{k=1}^r (\sum_{i=1}^n \sum_{j=1}^m x_{ijk})^2 / (n*m) - C \\ SS_e &= SS_T - SS_t - SS_u \end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^r x_{ijk})^2 / (n*m*r)$. The total degrees of freedom, the degrees of freedom of factor A, the degrees of freedom of factor B, the degrees of freedom of treatments' combination, the degrees of freedom of interactions between factor A and factor B, and the degrees of freedom of blocks, and degrees of freedom of errors are

$$\begin{aligned} df_T &= n*m*r - 1 \\ df_a &= n - 1 \\ df_b &= m - 1 \\ df_t &= n*m - 1 \\ df_{ab} &= (n-1)*(m-1) \\ df_u &= r - 1 \\ df_e &= (n*m-1)*(r-1) \end{aligned}$$

Calculate F -values for factor A, factor B, the interactions between factor A and factor B, and blocks as follows

$$\begin{aligned} F_a &= (SS_a/df_a) / (SS_e/df_e) \\ F_b &= (SS_b/df_b) / (SS_e/df_e) \\ F_{ab} &= (SS_{ab}/df_{ab}) / (SS_e/df_e) \\ F_u &= (SS_u/df_u) / (SS_e/df_e) \end{aligned}$$

(1) Effect size testing

Calculate the effect size measure, Cohen's η^2 (Cohen, 1988; Li, 2021; Zhang, 2022c), for factor A, factor B, and interactions between factor A and factor B

$$\begin{aligned} \eta_a^2 &= F_a df_a / (F_a df_a + df_e) \\ \eta_b^2 &= F_b df_b / (F_b df_b + df_e) \\ \eta_{ab}^2 &= F_{ab} df_{ab} / (F_{ab} df_{ab} + df_e) \end{aligned}$$

η_a^2 , η_b^2 , or $\eta_{ab}^2 < 0.10$, no effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \in [0.10, 0.30)$, small effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \in [0.30, 0.50)$, intermediate effect; η_a^2 , η_b^2 , or $\eta_{ab}^2 \geq 0.50$, large effect.

If there is at least an intermediate interaction effect (or small interaction effect, or large interaction effect, which is dependent upon research's requirement), i.e., $\eta_{ab}^2 \geq 0.30$, ANOVA should be terminated. Otherwise, continue for next steps.

If there is at least an intermediate factor effect (or small factor effect, or large factor effect, which is dependent upon research's requirement), i.e., $\eta_a^2 \geq 0.30$ (or $\eta_b^2 \geq 0.30$), use the effect size measure, Cohen's d , to compare between-mean differences of the factor A (or B) (Cohen, 1988; Zhang, 2023). For a pair of levels i and j ($i, j = 1, 2, \dots, n$; $i \neq j$) of factor A (or B; $i, j = 1, 2, \dots, m$; $i \neq j$), Cohen's d is

$$d = (m_j - m_i) / ((s_i^2 + s_j^2) / 2)^{0.5}$$

where m_i and m_j are the mean of the two levels i and j , respectively, and s_i and s_j are the standard deviation of the two levels i and j . $|d| < 0.20$, no level effect; $|d| \in [0.20, 0.50)$, small level effect; $|d| \in [0.50, 0.80)$, intermediate level effect; $|d| \geq 0.80$, large level effect.

(2) Significance testing

There are significant differences among treatments (levels) of factor A, if $F > F_\alpha(df_a, df_e)$; there are significant differences among treatments (levels) of factor B, if $F_b > F_\alpha(df_b, df_e)$; there are significant differences among blocks, if $F_u > F(df_u, df_e)$; there are significant differences among the interactions between factor A and factor B, if $F_{ab} > F_\alpha(df_{ab}, df_e)$. $\alpha = 0.001, 0.0001$, etc.

If $F_{ab} \leq F_\alpha(df_{ab}, df_e)$, use LSD method to compare between-mean differences. Calculate

$$S_a = (2 * (SS_e / df_e) / (m * r))^{1/2}$$

$$S_b = (2 * (SS_e / df_e) / (n * r))^{1/2}$$

And the corresponding LSDs are

$$LSD_a = t_\alpha(df_e) * S_a$$

$$LSD_b = t_\alpha(df_e) * S_b$$

Calculate the difference of pairwise means above and if it is greater than the corresponding LSD_a , then the difference of means are significant.

2.3 Multi-Factor ANOVA

2.3.1 Split Design

Assuming that blocks are randomly designed, and main treatments and sub-treatments are fixed. The data matrix is (x_{ijk}) , $i = 1, \dots, n$; $j = 1, 2, \dots, h$; $k = 1, 2, \dots, m$, where n is the number of main treatments, h is the number of sub-treatments, and m is the number of blocks. Calculate the total sum of squares (SS_T), sum of squares of main-blocks (SS_m), the sum of squares of blocks (SS_r), the sum of squares of main treatments (SS_a), the sum of squares of errors of main blocks (SS_{ea}), the sum of squares of treatments' combination (SS_l), the sum of squares of sub-treatments (SS_b), the sum of squares of the interactions between main treatments and sub-treatments (SS_{ab}), and the sum of squares of errors of sub-blocks (SS_{eb}) as follows

$$\begin{aligned}
SS_T &= \sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk}^2 - C \\
SS_m &= \sum_{i=1}^n \sum_{k=1}^m (\sum_{j=1}^h x_{ijk})^2 / h - C \\
SS_r &= \sum_{k=1}^m (\sum_{i=1}^n \sum_{j=1}^h x_{ijk})^2 / (n * h) - C \\
SS_a &= \sum_{i=1}^n (\sum_{k=1}^m \sum_{j=1}^h x_{ijk})^2 / (m * h) - C \\
SS_{ea} &= SS_m - SS_r - SS_a \\
SS_t &= \sum_{i=1}^n \sum_{j=1}^h (\sum_{k=1}^m x_{ijk})^2 / m - C \\
SS_b &= \sum_{j=1}^h (\sum_{i=1}^n \sum_{k=1}^m x_{ijk})^2 / (n * m) - C \\
SS_{ab} &= SS_t - SS_a - SS_b \\
SS_{eb} &= SS_T - SS_r - SS_t - SS_{ea}
\end{aligned}$$

where $C = (\sum_{i=1}^n \sum_{j=1}^h \sum_{k=1}^m x_{ijk})^2 / (n * h * m)$. The total degrees of freedom, the degrees of freedom of main-blocks, the degrees of freedom of blocks, the degrees of freedom of main treatments, the degrees of freedom of main blocks, the degrees of freedom of treatments' combination, the degrees of freedom of sub-treatments, the degrees of freedom of the interactions between main treatments and sub-treatments, and the degrees of freedom of sub-blocks are

$$\begin{aligned}
df_t &= n * h * m - 1 \\
df_m &= n * m - 1 \\
df_r &= m - 1 \\
df_a &= n - 1 \\
df_{ea} &= df_m - df_r - df_a \\
df_t &= n * h - 1 \\
df_b &= h - 1 \\
df_{ab} &= (n - 1) * (h - 1) \\
df_{eb} &= n * (h - 1) * (m - 1)
\end{aligned}$$

Calculate F -values for main treatments, sub-treatments, and the interactions between main treatments and sub-treatments as follows

$$\begin{aligned}
F_a &= (SS_a / df_a) / (SS_{ea} / df_{ea}) \\
F_b &= (SS_b / df_b) / (SS_{eb} / df_{eb}) \\
F_{ab} &= (SS_{ab} / df_{ab}) / (SS_{eb} / df_{eb})
\end{aligned}$$

For significance testing, there are significant differences among main treatments, if $F > F_{\alpha}(df_a, df_{ea})$; there are significant differences among sub-treatments, if $F_b > F_{\alpha}(df_b, df_{eb})$; there are significant differences among the interactions between main treatments and sub-treatments, if $F_{ab} > F_{\alpha}(df_{ab}, df_{eb})$. $\alpha = 0.001, 0.0001$, etc.

If $F_{ab} \leq F_{\alpha}(df_{ab}, df_{eb})$, use LSD method to compare between-mean differences. Calculate

$$\begin{aligned}
S_a &= (2 * (SS_{ea} / df_{ea}) / (h * m))^{1/2} \\
S_b &= (2 * (SS_{eb} / df_{eb}) / (n * m))^{1/2} \\
S_{ab} &= (2 * (SS_{eb} / df_{eb} * (h - 1) + SS_{ea} / df_{ea}) / (h * m))^{1/2}
\end{aligned}$$

And the corresponding LSDs are

$$LSD_{\alpha} = t_{\alpha}(df_{ea}) * S_a$$

$$LSD_{\alpha} = t_{\alpha}(df_{eb}) * S_b$$

$$LSD_{\alpha} = S_{ab} * ((h-1) * SS_{eb} / df_{eb} * t_{\alpha}(df_{eb}) + SS_{ea} / df_{ea} * t_{\alpha}(df_{ea})) / ((h-1) * SS_{eb} / df_{eb} + SS_{ea} / df_{ea})$$

Calculate the difference of pairwise means above and if it is greater than the corresponding LSD_{α} , then the difference of means are significant.

2.3.2 Orthogonal Design

Assuming that blocks are randomly designed, the data matrix is (x_{ij}) , $i=1, \dots, n$; $j=1, 2, \dots, w+(m-w)+r$, where n is the number of treatments, w is the number of factors, and m is the number of blank columns, and r the number of replications. x_{ij} are the levels of factors, $i=1, \dots, n$; $j=1, 2, \dots, w+(m-w)$; x_{ij} are experimental results, $i=1, \dots, n$; $j=w+(m-w)+1, \dots, w+(m-w)+r$. Assuming that q is the number of levels of the factor. First, calculate

$$s_i = \sum_{j=m+1}^{m+r} x_{ij}, \quad i=1, \dots, n$$

$$kk_{ij} = \sum_{k=1}^n s_k, \quad i=1, \dots, q; j=1, 2, \dots, m; x_{kj}=i$$

$$p_{ij} = kk_{ij} / (w * r)$$

$$r_j = \max p_{ij} - \min p_{ij}, \quad i=1, \dots, q; j=1, 2, \dots, w$$

Calculate the total sum of squares (SS_T), sum of squares of blocks (SS_r), the sum of squares of factors (SS_j), the sum of squares of columns (SS_c), and the sum of squares of errors (SS_e) as follows

$$SS_T = \sum_{i=1}^n \sum_{j=m+1}^{m+r} x_{ij}^2 - C$$

$$SS_r = \sum_{j=m+1}^{m+r} \sum_{i=1}^n (\sum_{j=1}^h x_{ij})^2 / n - C$$

$$SS_j = \sum_{i=1}^q kk_{ij}^2 / (r * q) - C, \quad j=1, 2, \dots, w;$$

$$SS_c = \sum_{j=w+1}^m \sum_{i=1}^q kk_{ij}^2 / (r * q) - C$$

$$SS_e = SS_T - SS_r - SS_c - \sum_{j=1}^w SS_j$$

where $C = (\sum_{i=1}^n s_i)^2 / (n * r)$. The total degrees of freedom, degrees of freedom of blocks, the degrees of freedom of factors, the degrees of freedom of columns, and the degrees of freedom of errors are

$$df_T = n * r - 1$$

$$df_r = r - 1$$

$$df_j = q - 1$$

$$df_c = (m - w) * (q - 1)$$

$$df_e = df_T - df_r - n + 1$$

For significance testing, calculate F -values for blocks, factors, and empty columns as follows

$$F_r = (SS_r / df_r) / (SS_e / df_e)$$

$$F_j = (SS_j / df_j) / (SS_e / df_e)$$

$$F_c = (SS_c / df_c) / (SS_e / df_e)$$

There are significant differences among blocks, if $F_r > F_{\alpha}(df_r, df_e)$; there are significant differences among factors, if $F_j > F_{\alpha}(df_j, df_e)$, and there are significant differences among blank columns, if $F_c > F_{\alpha}(df_c, df_e)$. $\alpha=0.001$, 0.0001, etc.

2.4 Effect size testing and significance testing

In addition to the adoption of more stringent p -values in significance testing, effect size testing in present study is a methodology of ANOVA in the paradigm of new statistics (effect size testing for Split Design and Orthogonal Design will be created in the future). To avoid the shortcomings of p -value based significance testing in traditional statistics, the results of effect size testing should be the prioritized consideration in ANOVA, supplemented by the results of significance testing, or at least the two results should be verified mutually. More ideally, all ANOVA results are suggested being verified by other professional knowledge and experimental facts in order to draw the most reliable conclusions.

3 Description of Data Matrix

For offline tool, the data matrix used in ANOVA is saved in the space delimited plain text file. The demonstration data matrices for the ANOVA methods, Randomized Complete Design (ANOVA-1), Randomized Block Design (ANOVA-2), Randomized Block Design with Subsamples (ANOVA-3), Nested Design (ANOVA-4), Two-Factor Classification (ANOVA-5), Two-Factor Randomized Complete Block Design (ANOVA-6), Split Design (ANOVA-7), and Orthogonal Design (ANOVA-8) in present study are indicated in Fig. 1.

The data formats are described as follows (Fig. 1)

Randomized Complete Design (ANOVA-1): No. treatments (total rows) = 6, No. replications (total columns) = 5.

Randomized Block Design (ANOVA-2): No. treatments (total rows) = 6, No. replications (total columns) = 5.

Randomized Block Design with Subsamples (ANOVA-3): No. treatments (total matrices) = 3, No. blocks (total columns) = 5, No. sub-samples (total rows in a matrix) = 6.

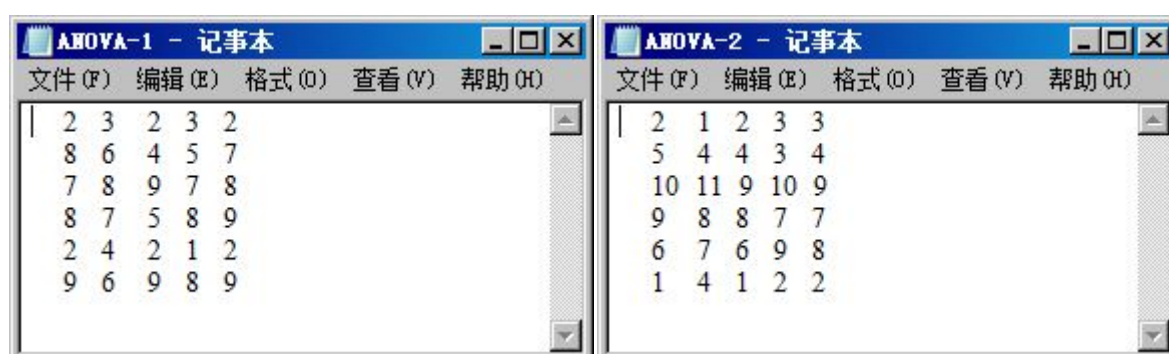
Nested Design (ANOVA-4): No. treatments (total matrices) = 3, No. blocks (total rows in a matrix) = 6, No. sub-samples (total columns) = 5.

Two-Factor Classification (ANOVA-5): No. levels of factor A (total matrices) = 3, No. levels of factor B (total columns) = 5, No. replications (total rows in a matrix) = 6.

Two-Factor Randomized Complete Block Design (ANOVA-6): No. levels of factor A (total matrices) = 3, No. levels of factor B (total rows in a matrix) = 6, No. replications (total columns) = 5.

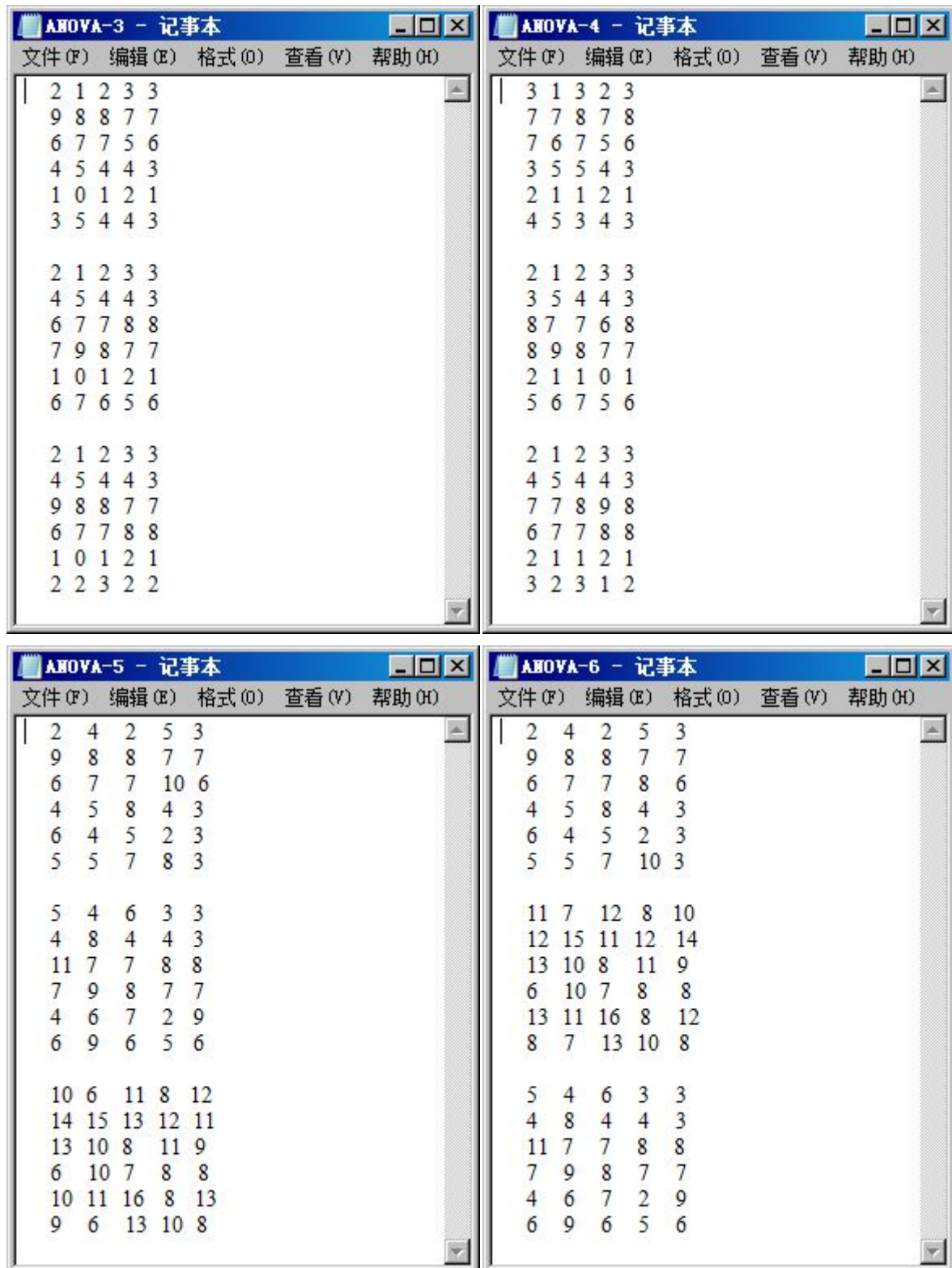
Split Design (ANOVA-7): No. main treatments (total matrices) = 3, No. sub-treatments (total rows in a matrix) = 6, No. blocks (total columns) = 5.

Orthogonal Design (ANOVA-8): No. treatments (total rows) = 9, No. factors (IDs 1, 2 and 3 in the first 4 columns) = 3, No. factor levels (the first 3 columns) = 3, No. columns (the first 4 columns) = 4, No. replications (the last 2 columns) = 2.



2	3	2	3	2
8	6	4	5	7
7	8	9	7	8
8	7	5	8	9
2	4	2	1	2
9	6	9	8	9

2	1	2	3	3
5	4	4	3	4
10	11	9	10	9
9	8	8	7	7
6	7	6	9	8
1	4	1	2	2



ANOVA-7 - 记事本

文件 (F) 编辑 (E) 格式 (O) 查看 (V) 帮助 (H)

4	5	4	4	3
1	0	1	2	1
2	1	2	3	3
9	8	8	7	7
6	7	7	5	6
3	5	4	4	3
6	7	7	8	8
7	9	8	7	7
2	1	2	3	3
4	5	4	4	3
1	0	1	2	1
6	7	6	5	6
6	7	7	8	8
1	0	1	2	1
2	1	2	3	3
4	5	4	4	3
9	8	8	7	7
2	2	3	2	2

ANOVA-8 - 记事本

文件 (F) 编辑 (E) 格式 (O) 查看 (V) 帮助 (H)

1	1	1	1	195	189
1	2	2	2	161	178
1	3	3	3	196	186
2	1	2	3	160	155
2	2	3	1	176	176
2	3	1	2	176	172
3	1	3	2	167	177
3	2	1	3	171	181
3	3	2	1	159	160

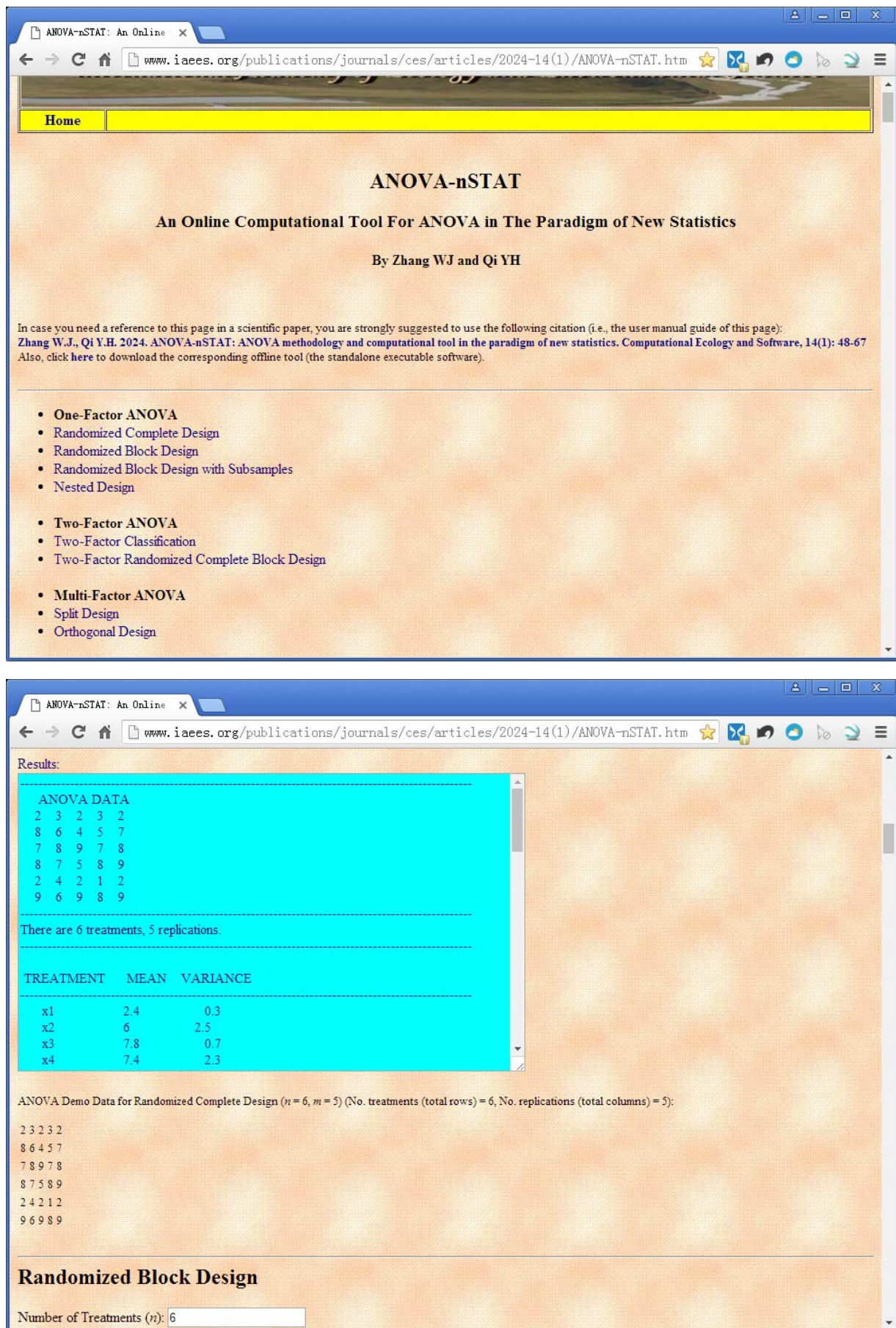
Fig. 1 Data formats of various ANOVA methods.

4 Online and Offline Computational Tools

4.1 Online tool

The online tool, ANOVA-nSTAT.htm, for the ANOVA above is a web page that can be used on various web browsers (Fig. 2). Therefore, it is a platform independent online computational tool. It was developed based on JavaScript:

[http://www.iaees.org/publications/journals/ces/articles/2024-14\(1\)/ANOVA-nSTAT.htm](http://www.iaees.org/publications/journals/ces/articles/2024-14(1)/ANOVA-nSTAT.htm)



ANOVA-nSTAT: An Online

www.iaees.org/publications/journals/ces/articles/2024-14(1)/ANOVA-nSTAT.htm

Home

ANOVA-nSTAT

An Online Computational Tool For ANOVA in The Paradigm of New Statistics

By Zhang WJ and Qi YH

In case you need a reference to this page in a scientific paper, you are strongly suggested to use the following citation (i.e., the user manual guide of this page):
 Zhang W.J., Qi Y.H. 2024. ANOVA-nSTAT: ANOVA methodology and computational tool in the paradigm of new statistics. *Computational Ecology and Software*, 14(1): 48-67
 Also, click [here](#) to download the corresponding offline tool (the standalone executable software).

- **One-Factor ANOVA**
 - Randomized Complete Design
 - Randomized Block Design
 - Randomized Block Design with Subsamples
 - Nested Design
- **Two-Factor ANOVA**
 - Two-Factor Classification
 - Two-Factor Randomized Complete Block Design
- **Multi-Factor ANOVA**
 - Split Design
 - Orthogonal Design

Results:

ANOVA DATA

```
2 3 2 3 2
8 6 4 5 7
7 8 9 7 8
8 7 5 8 9
2 4 2 1 2
9 6 9 8 9
```

There are 6 treatments, 5 replications.

TREATMENT	MEAN	VARIANCE
x1	2.4	0.3
x2	6	2.5
x3	7.8	0.7
x4	7.4	2.3

ANOVA Demo Data for Randomized Complete Design ($n = 6, m = 5$) (No. treatments (total rows) = 6, No. replications (total columns) = 5):

```
2 3 2 3 2
8 6 4 5 7
7 8 9 7 8
8 7 5 8 9
2 4 2 1 2
9 6 9 8 9
```

Randomized Block Design

Number of Treatments (n):

Fig. 2 The web page of the online tool, ANOVA-nSTAT.htm

4.2 Offline tool

The offline tool, ANOVA-nSTAT.exe, for the ANOVA above is the standalone executable software that can be used on various Windows operational systems (Fig. 3, 4). It was developed using Delphi development tool. Included in a supplementary material package, the offline tool can be freely downloaded at:

[http://www.iaees.org/publications/journals/ces/articles/2024-14\(1\)/4-Zhang-Abstract.asp](http://www.iaees.org/publications/journals/ces/articles/2024-14(1)/4-Zhang-Abstract.asp)

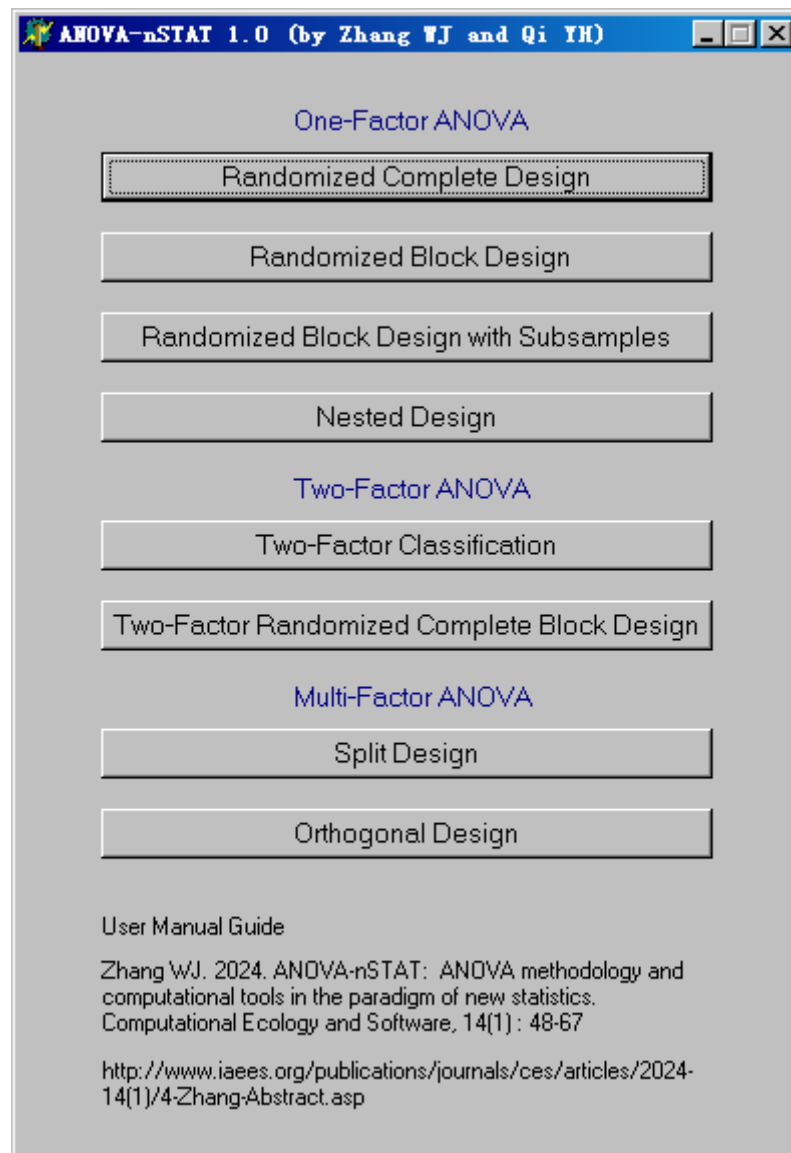


Fig. 3 Main window of offline tool, ANOVA-nSTAT.exe

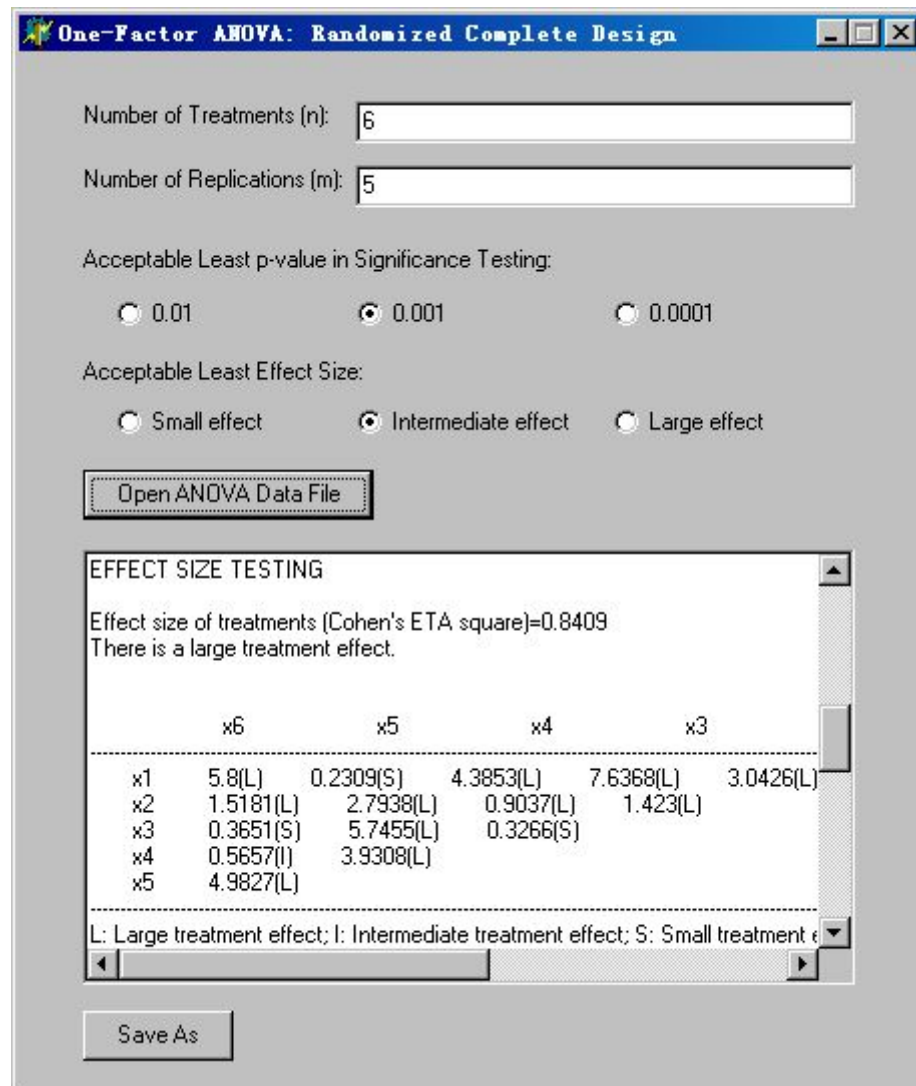


Fig. 4 Main window of Randomized Complete Design of offline tool, ANOVA-nSTAT.exe

5 An Example

We use the demonstration data, ANOVA-1, for the analysis of Randomized Complete Design. The computational results of offline tool are as follows

ANOVA DATA

```

2 3 2 3 2
8 6 4 5 7
7 8 9 7 8
8 7 5 8 9
2 4 2 1 2
9 6 9 8 9

```

There are 6 treatments, 5 replications.

TREATMENT	MEAN	VARIANCE
x1	2.4	0.3
x2	6	2.5
x3	7.8	0.7
x4	7.4	2.3
x5	2.2	1.2
x6	8.2	1.7

EFFECT SIZE TESTING

Effect size of treatments (Cohen's η^2)=0.8409

There is a large treatment effect.

	x6	x5	x4	x3	x2
x1	5.8(L)	0.2309(S)	4.3853(L)	7.6368(L)	3.0426(L)
x2	1.5181(L)	2.7938(L)	0.9037(L)	1.423(L)	
x3	0.3651(S)	5.7455(L)	0.3266(S)		
x4	0.5657(I)	3.9308(L)			
x5	4.9827(L)				

L: Large treatment effect; I: Intermediate treatment effect; S: Small treatment effect; N: No treatment effect.
Effect size measure: Cohen's d.

SIGNIFICANCE TESTING

TABLE OF VARIANCE ANALYSIS

SOURCE	SS	DF	MS	F
TREATMENTS	183.8667	5	36.7733	25.3609**
ERROR	34.8	24	1.45	

**.: p=0.0001; *: p=0.001

LSD TESTS

	x5	x1	x2	x4	x3
x6	6**	5.8**	2.2	0.8	0.4
x3	5.6**	5.4**	1.8	0.4	
x4	5.2**	5**	1.4		
x2	3.8*	3.6*			
x1	0.2				

** : $p=0.0001$; * : $p=0.001$.

For effect size testing, we can find that treatment effect is large. There is an intermediate treatment effect between the treatments 4 and 6. There are small treatment effects between the treatments 3 and 4 and between the treatments 3 and 6. There are large treatment effects for other paired treatments. In significance testing, the overall treatment effect is significant at $p=0.0001$ (confidence degree: 99.99%). However, there are not significant differences between the treatments 3 and 4 and the treatments 3 and 6. In these cases, we deeply confirm the common results of overall treatment effects of both testing, and tend to accept the results of effect size testing for the comparisons between the treatments 3 and 4 and between the treatments 3 and 6, if no more valuable information is available.

References

- Amrhein V, Greenland S, McShane B. 2019. Scientists rise up against statistical significance. *Nature*, 567: 305-307. <https://doi.org/10.1038/d41586-019-00857-9>
- Bergstrom CT, West JD. 2021. Manipulated P-values: Mathematical nonsense in scientific papers. https://www.laitimes.com/en/article/3km6i_41b77.html. Accessed 2022-4-23
- Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Erlbaum, Hillsdale, NJ, USA. <https://doi.org/10.4324/9780203771587>
- Cohen BH. 2008. *Explaining Psychological Statistics* (3rd ed). John Wiley and Sons, New York, USA
- Cumming G. 2013. The new statistics: Why and how. *Psychological Science*, 25: 7-29. <https://doi.org/10.1177/0956797613504966>
- Ellis PD. 2010. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press, Cambridge, UK. <https://doi.org/10.1017/CBO9780511761676>
- Errington TM, Mathur M, Soderberg CK, et al. 2021. Investigating the replicability of preclinical cancer biology. *eLife*, 10: e71601. <https://elifesciences.org/articles/71601>
- Fisher RA. 1935. *The Design of Experiments*. Oliver and Boyd, Edinburgh and London, UK. <http://www.medicine.mcgill.ca/epidemiology/hanley/tmp/Mean-Quantile/DesignofExperimentsCh-III.pdf>
- Huang HN. 2023. Statistics reform: practitioner's perspective. DOI: 10.13140/RG.2.2.31799.50084
- Huang HN. 2021a. Statistics reform: challenges and opportunities. *ScienceNet*. <https://blog.sciencenet.cn/blog-3427112-1318043.html>. Accessed 2021-12-26

- Huang HN. 2021b. What are the most misunderstood and misleading concepts or theories in statistics textbooks today? ScienceNet. <https://blog.sciencenet.cn/blog-3427112-1269013.html>. Accessed 2021-1-26
- Ioannidis JPA. 2005. Why most published research findings are false. Plos Medicine, <https://doi.org/10.1371/journal.pmed.0020124>
- Kafdar K. 2021. Editorial: Statistical significance, p -values, and replicability. The Annals of Applied Statistics. <https://doi.org/10.1214/21-AOAS1500>
- Lenhard W, Lenhard A. 2016. Computation of effect sizes. https://www.psychometrica.de/effect_size.html. Psychometrica. doi: 10.13140/RG.2.2.17823.92329
- Li HH. 2021. p -values are too sensitive, and the effect size is long and good to save. ScienceNet. <http://blog.sciencenet.cn/blog-2619783-1286084.html>. Accessed 2021-5-11
- McShane BB, David Gal D. 2017. Statistical significance and the dichotomization of evidence. Journal of the American Statistical Association, 112(519): 885-895. <https://doi.org/10.1080/01621459.2017.1289846>
- Nature Editorial. 2021. Replicating scientific results is tough — but essential. Nature, 600: 359-360. <https://doi.org/10.1038/d41586-021-03736-4>
- Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. Science, 349(6251). <https://doi.org/10.1126/science.aac4716>
- Sellke T, Bayarri MJ, Berger JO. 2001. Calibration of p values for testing precise null hypotheses. The American Statistician, 55(1): 62-71. <https://doi.org/10.1198/000313001300339950>
- Sun XJ. 2016. Improper use of p -values is a nonsense. ScienceNet. <http://blog.sciencenet.cn/blog-41174-961169.html>. Accessed 2016-4-5
- Tong C. 2019. Statistical Inference Enables Bad Science; Statistical Thinking Enables Good Science. The American Statistician, 73(S1): 246-261. <https://doi.org/10.1080/00031305.2018.1518264>
- Vrieze JD. 2021. Landmark research integrity survey finds questionable practices are surprisingly common. Science. <https://doi.org/10.1126/science.abk3508>
- Wasserstein RL, Schirm AL, Lazar NA, 2019. Editorial: Moving to a world beyond “ $p < 0.05$ ”. The American Statistician, 79: 1-19. <https://doi.org/10.1080/00031305.2019.1583913>
- Yates F. 1951. The influence of statistical methods for research workers on the development of the science of statistics. Journal of the American Statistical Association, 46: 19-34. <https://doi.org/10.2307/2280090>
- Zhang WJ. 2022a. Confidence intervals: Concepts, fallacies, criticisms, solutions and beyond. Network Biology, 12(3): 97-115. [http://www.iaees.org/publications/journals/nb/articles/2022-12\(3\)/confidence-intervals-fallacies-criticisms-solutions.pdf](http://www.iaees.org/publications/journals/nb/articles/2022-12(3)/confidence-intervals-fallacies-criticisms-solutions.pdf)
- Zhang WJ. 2022b. Dilemma of t -tests: Retaining or discarding choice and solutions. Computational Ecology and Software, 12(4): 181-194. [http://www.iaees.org/publications/journals/ces/articles/2022-12\(4\)/dilemma-of-t-tests.pdf](http://www.iaees.org/publications/journals/ces/articles/2022-12(4)/dilemma-of-t-tests.pdf)
- Zhang WJ. 2022c. p -value based statistical significance tests: Concepts, misuses, critiques, solutions and beyond. Computational Ecology and Software, 12(3): 80-122. [http://www.iaees.org/publications/journals/ces/articles/2022-12\(3\)/p-value-based-statistical-significance-tests.pdf](http://www.iaees.org/publications/journals/ces/articles/2022-12(3)/p-value-based-statistical-significance-tests.pdf)
- Zhang WJ. 2023. A desktop calculator for effect sizes: Towards the new statistics. Computational Ecology and Software, 13(4): 136-181 [http://www.iaees.org/publications/journals/ces/articles/2023-13\(4\)/4-Zhang-Abstract.asp](http://www.iaees.org/publications/journals/ces/articles/2023-13(4)/4-Zhang-Abstract.asp)
- Zhang WJ, Qi YH. 2005. Algorithms integration and network implementation of eight univariate variance analyses. Computer Applications and Software, 22(3): 122-125