

# combinIT: An R Package for Combining Interaction Tests for Unreplicated Two-Way Tables

by Mahmood Kharrati-Kopaei, Zahra Shenavari, Hossein Haghbin

**Abstract** Several new tests have been proposed for testing interaction in unreplicated two-way analysis of variance models. Unfortunately, each test is powerful for detecting a pattern of interaction. Therefore, it is reasonable to combine multiple interaction tests to increase the power of detection for significant interactions. We introduce the package **combinIT** that provides researchers the results of six existing recommended interaction tests, including: the value of test statistics, exact Monte Carlo p-values, approximated or adjusted p-values, the results of four combined tests and explanations of interaction types if the discussed tests are significant. The software **combinIT** is a more comprehensive R package in comparison with the two existing packages. In addition, the software is executed quickly to obtain the exact Monte Carlo p-values, even for large Monte Carlo runs, in contrast to existing packages.

## 1 Introduction

Suppose that there are two factors  $A$  and  $B$  with  $a$  and  $b$  levels, respectively. To investigate the effect of the factors on a response variable  $y$ , an unreplicated two-way analysis of variance (ANOVA) model is sometimes used to reduce the experiment cost. Formally, this model is

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij}, \quad i = 1, \dots, a; \quad j = 1, \dots, b, \quad (1)$$

where  $y_{ij}$  denotes the  $i$ -th and the  $j$ -th observation of the response variable,  $\mu$  is the grand mean,  $\alpha_i$  and  $\beta_j$  denote the main effects of the factors,  $\gamma_{ij}$ s are interaction terms, and  $\epsilon_{ij}$ s are random errors which are independent and identically normal variables with mean zero and variance  $\sigma^2$ . As noted by [Franck and Osborne \(2016\)](#), the unreplicated two-way ANOVA models arise in the context of (a) completely randomized two-factor experiments, (b) randomized complete block experiments that block on a source of nuisance variability, and (c) observational studies with two factors. The main problem in an unreplicated two-way ANOVA model is that all observations are used to estimate the model parameters and no observations remain to estimate  $\sigma^2$ ; see [Shenavari and Kharrati-Kopaei \(2018\)](#) and [Kharrati-Kopaei and Sadooghi-Alvandi \(2007\)](#). Therefore, neither F-test nor t-test can be used to make inferences on the effects (including the interaction effect). The textbook approach is to assume that the interaction terms are zero (i.e. the model is additive). In this case,  $(a - 1)(b - 1)$  degrees of freedom are left to estimate  $\sigma^2$  and the usual F-tests can be used to make inferences about the main effects. However, if the interaction effects are present (i.e. the model is non-additive), the main effects may be masked by the interaction effect; see [Montgomery \(2017\)](#). Therefore, it is important to implement some method to test if there exists a significant interaction in an unreplicated ANOVA model.

The existing approaches to test interaction in unreplicated two-way ANOVA models can be categorized into two main groups: functional-based-form (FBF) and non-functional-based-form (NFBF). In FBF approaches, a functional form is assumed for the interaction terms. This group includes interaction tests proposed by [Tukey \(1949\)](#), [Mandel \(1961, 1971\)](#), [Johnson and Graybill \(1972\)](#), and [Corsten and Eijnsbergen \(1972, 1974\)](#); see [Milliken and Johnson \(1989, pp 2–63\)](#) for a discussion of some of these interaction tests. [Simecek and Simeckova \(2013\)](#) also modified Tukey's test. It is known that these tests are powerful for detecting interaction only when the specified functional form for the interaction terms is appropriate; see [Boik \(1993a\)](#), [Alin and Kurt \(2006\)](#), [Simecek and Simeckova \(2013\)](#), and [Shenavari and Kharrati-Kopaei \(2018\)](#). In NFBF approaches, a specific functional form is not assumed for the interaction terms. The NFBF group includes interaction tests proposed by [Milliken and Rasmuson \(1977\)](#), [Tusell \(1990\)](#), [Boik \(1993a\)](#), [Piepho \(1994\)](#), [Kharrati-Kopaei and Sadooghi-Alvandi \(2007\)](#), [Franck et al. \(2013\)](#), [Malik et al. \(2016\)](#) and [Kharrati-Kopaei and Miller \(2016\)](#). [Kharrati-Kopaei and Sadooghi-Alvandi \(2007\)](#) formally showed that there does not exist the best interaction test in an unreplicated two-way ANOVA model in the sense that there is no test that can detect all patterns of interaction with high power. This means that each test is powerful for detecting certain types of interaction. Therefore, it is meaningful to combine multiple interaction tests to provide researchers with a testing approach that leverages many existing methods to detect different patterns of non-additivity. We note that the interaction tests are dependent and the main problem is how to combine multiple dependent tests into a single test procedure for testing additivity such that it controls the Type I error rate and has an acceptable power. [Shenavari and Kharrati-Kopaei \(2018\)](#) evaluated the performances

of different combination methods and ultimately recommended combining the interaction tests by the Bonferroni method, a Beta (Sidak) approximation, the Jacobi polynomial expansion, and the Gaussian copula method. These methods are abbreviated as Bon, Sidak, JPE, and GC, respectively, throughout this paper.

Despite the application of unreplicated two-way ANOVA models in industry, biology, agriculture, and medicine, the newly proposed interaction tests have not been discussed in statistical packages; as similarly noted by Franck and Osborne (2016). Recently, Osborne et al. (2016) released the package **hiddenf** which is available from the Comprehensive R Archive Network (CRAN). As the name of the package **hiddenf** implies, this package mainly focuses on detecting a hidden non-additivity structure proposed by Franck et al. (2013). This package also reports the p-values of tests for non-additivity developed by Tukey (1949), Mandel (1961), Kharrati-Kopaei and Sadooghi-Alvandi (2007), and Malik et al. (2016). However, this package neither provides the result of NFBF interaction tests proposed in Boik (1993a), Piepho (1994), and Kharrati-Kopaei and Miller (2016) nor the result of the combined interaction test. In addition, this package uses a Monte Carlo procedure to report the p-value of Malik et al. (2016); however, the applied procedure is rather time-consuming. Further, the package **hiddenf** provides only adjusted Bonferroni p-values of the tests proposed by Kharrati-Kopaei and Sadooghi-Alvandi (2007) and Franck et al. (2013) instead of the exact Monte Carlo p-values. Although Franck and Osborne (2016) mentioned that the Bonferroni adjustment is not overly conservative for  $\alpha \leq 7$ , one might be interested in knowing the exact p-values. We note that Simeckova et al. (2014) also developed the **additivityTests** package that provides test statistics, critical values, and binary reject/fail-to-reject decisions for tests proposed by Tukey (1949), Mandel (1961), Boik (1993a), Tusell (1990), Johnson and Graybill (1972), and Simecek and Simeckova (2013). However, the **additivityTests** package does not provide the p-values of tests; see Franck and Osborne (2016).

In this paper, we introduce the R package **combinIT**, which is available from CRAN (Shenavari et al., 2022). This package reports both exact Monte Carlo and adjusted or approximate (if available) p-values of the six NFBF interaction tests developed by Boik (1993a), Piepho (1994), Kharrati-Kopaei and Sadooghi-Alvandi (2007), Franck et al. (2013), Malik et al. (2016), and Kharrati-Kopaei and Miller (2016). We use abbreviations Boik, Piepho, KKSA, Franck, Malik, and KKM, respectively, for these NFBF tests. In addition, this package provides the results of four combined interaction tests that are based on the Bon, Sidak, JPE, and GC methods. Furthermore, if a significant interaction is detected by a combined test, the package **combinIT** gives some explanations of the interaction type or pattern. Note that FBF tests have not been considered in **combinIT** because Boik (1993b) showed that the Boik test is never less powerful and is sometimes much more powerful than the test proposed by Johnson and Graybill (1972), which can be regarded as an extension of FBF tests proposed by Tukey (1949) and Mandel (1961, 1971). In addition, Boik (1993b) compared the Boik test with the NFBF test developed by Tusell (1990) and recommended the Boik test. Furthermore, the Piepho test is a modified version of the NFBF test proposed by Milliken and Rasmuson (1977) hence this test has not been considered in **combinIT**, either. In summary, the **combinIT** package reports the results of all recommended existing interaction tests in unreplicated two-way ANOVA models and, in this view, can be regarded as a more comprehensive package than **hiddenf** for testing interaction. In terms of code execution speed, nearly 25% of the code has been written in C++; see <https://github.com/hagbinh/combinIT>. We used the **Rcpp** package (Eddelbuettel and François, 2011; Eddelbuettel, 2013; Eddelbuettel and Balamuta, 2018) for writing some parts of the codes in C++. Thus, Monte Carlo simulations for calculating the p-value of the Malik test are not as time-consuming as those in the **hiddenf** package. In addition, using the **Rcpp** package allows us to calculate the exact Monte Carlo p-values of the Franck and KKSA tests in a reasonable time, in contrast to the **hiddenf** package, which provides only adjusted Bonferroni p-values; see Franck and Osborne (2016).

The rest of the paper is organized as follows. In the next section, the six NFBF interaction tests and the four combination methods that are utilized in the **combinIT** package are briefly reviewed. Then, technical details of the **combinIT** package are provided. After that, use of the package **combinIT** is illustrated with an example. The execution speed of the code is also discussed in this section. Finally, the performances of the combined interaction tests in terms of controlling the Type I error rate are discussed via a simulation study by using the **hiddenf** and **combinIT** packages.

## 2 Six NFBF interaction tests and four combination methods

This section has two subsections. We first review the Boik, Piepho, KKSA, Franck, Malik, and KKM interaction tests. We also discuss what patterns of interactions these tests are powerful for detecting. We then review the Bon, Beta, JPE, and GC combination methods. These tests and combination methods are utilized in the **combinIT** package. Throughout this section, the null hypothesis of interest is  $H_0$ : There is no interaction or, equivalently,  $H_0: \gamma_{ij} = 0$  for all  $i$  and  $j$  and  $r_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$ . (in usual notation) denote the residuals of the model 1 under  $H_0$ .

**Six NFBF interaction tests**

**Boik.** Let  $p = \min\{a - 1, b - 1\}$ ,  $q = \max\{a - 1, b - 1\}$  and  $\text{tr}(\mathbf{X})$  denote the trace of matrix  $\mathbf{X}$ . Boik (1993a) proposed a locally best invariant (LBI) test and  $H_0$  is rejected when

$$T_{\text{Boik}} = \text{tr}^2(\mathbf{R}'\mathbf{R}) / \left( \text{ptr} \left( (\mathbf{R}'\mathbf{R})^2 \right) \right)$$

is small, where  $\mathbf{R} = [r_{ij}]$  denotes the residual matrix of the model 1. Boik (1993a) provided the exact distribution of the test statistic under the null hypothesis  $H_0$  when  $p = 2$  and its asymptotic distribution when  $q$  is large. Note that the p-value of the Boik test is always 1 when  $p = 1$ . Note that the Boik (1993a) test is powerful for detecting interaction when the matrix of interaction terms  $[\gamma_{ij}]$  has small rank and one singular value dominates the remaining singular values; see Boik (1993b). For example, the Boik test is powerful for detecting the multiplicative form of interaction. The package **combinIT** provides the value of test statistic  $T_{\text{Boik}}$ , an asymptotic p-value, and an exact p-value of the Boik test based on a Monte Carlo simulation procedure.

**Piepho.** Piepho (1994) discussed and modified an interaction test procedure proposed by Milliken and Rasmuson (1977) by using the estimator of variance for each level of one factor; i.e.,

$$\hat{\sigma}_i^2 = \frac{a(a-1)W_i - \sum_{k=1}^a W_k}{(a-1)(a-2)(b-1)},$$

where  $W_i = \sum_{j=1}^b r_{ij}^2$ . Piepho (1994) proposed three test statistics. The **combinIT** package utilizes the third one, which detects a significant interaction when

$$T_{\text{Piepho}} = -(a-1)(a-2)(b-1) \log(U) / 2$$

is large, where  $U = 2a \sum_{i < j} \hat{\sigma}_i^2 \hat{\sigma}_j^2 / \left( (a-1) \left( \sum_{i=1}^a \hat{\sigma}_i^2 \right)^2 \right)$ . Piepho (1994) provided only an asymptotic chi-square distribution of  $T_{\text{Piepho}}$  under  $H_0$ . By construction, the Piepho test is powerful for detecting interactions when the estimators of variances are heterogeneous across the levels of one factor. The package **combinIT** provides the value of test statistic  $T_{\text{Piepho}}$ , an asymptotic chi-square p-value, and an exact p-value of the Piepho test based on a Monte Carlo simulation procedure.

**KKSA.** Suppose that  $a \geq b$  and  $b \geq 4$ . Split the data table into two sub-tables, obtained by putting  $a_1$  ( $2 \leq a_1 \leq a - 2$ ) rows in the first sub-table and the remaining  $a_2$  rows in the second sub-table ( $a_1 + a_2 = a$ ). The number of all possible divisions of the data table is  $\text{APD} = 2^{(a-1)} - a - 1$ . For the  $l$ -th division, let  $\text{RSS1}$  and  $\text{RSS2}$  denote the residual sum of squares for the two sub-tables. Let  $F_l^* = \max\{F_l, 1/F_l\}$  where  $F_l = (a_2 - 1) \text{RSS1} / ((a_1 - 1) \text{RSS2})$  and  $P_l$  denote the corresponding p-value. Kharrati-Kopaei and Sadooghi-Alvandi (2007) proposed  $\text{minP} = \min_{1 \leq l \leq \text{APD}} P_l$  as a test statistic and the hypothesis of no interaction is rejected for a small value of  $\text{minP}$ . The KKSA test is powerful for detecting interaction when the magnitude of interaction effects is heteroscedastic across the sub-tables of observations. The package **combinIT** provides the value of the test statistic and an exact p-value of the KKSA test based on a Monte Carlo simulation procedure. In addition, it provides an approximate p-value based on the Bonferroni adjustment. Note that the KKSA test is not applicable when both  $a$  and  $b$  are less than 4.

**Franck.** Franck et al. (2013) defined the hidden additivity structure as “the levels of one factor belong in two or more groups such that within each group the effects of the two factors are additive but the groups may interact with the ungrouped factor”. Based on this concept, Franck et al. (2013) proposed a test statistic by dividing the table of data into two sub-tables and developing an interaction F-test. Then, they considered all possible configurations of data and used the maximum of the interaction F-tests as a test statistic; the all-configurations and maximum-interaction F-test (ACMIF). The hypothesis of no interaction is rejected when ACMIF is large. It is clear that the Franck test is powerful when there is a hidden additivity structure in the data set. The package **combinIT** provides the value of the test statistic (ACMIF), an exact p-value of the Franck test based on a Monte Carlo simulation procedure, and an approximate p-value based on the Bonferroni adjustment.

**Malik.** Malik et al. (2016) used the idea that some cells may be involved in a significant interaction pattern and those cells might produce large positive or negative residuals. Therefore, Malik et al. (2016) proposed to partition the residuals into three clusters using a suitable clustering method like k-means clustering. The hypothesis of no interaction can be interpreted as the effect of the three clusters being equal. In more detail, Malik et al. (2016) proposed the model  $y_{ij} = \mu^* + \alpha_i^* + \beta_j^* + \zeta_{k(ij)} + \epsilon_{ij}$  where  $\zeta_k$  is the cluster effect for the  $k$ -th cluster,  $k = 1, 2, 3$ , and  $k(ij)$  denotes the cluster  $k$  to which the residual of the  $(i, j)$ -th cell is assigned by the clustering procedure. Malik et al. (2016) considered the hypothesis of no interaction as  $H_0 : \zeta_1 = \zeta_2 = \zeta_3$ . Let  $\text{SSE}(\text{cluster}) = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^3 \left( y_{ij} - \hat{\mu}^* - \hat{\alpha}_i^* - \hat{\beta}_j^* - \hat{\zeta}_{k(ij)} \right)^2$  where the hat notation denotes the estimates of effects. Malik et al. (2016) proposed the following test

statistic to test  $H_0$

$$T_{\text{Malik}} = \frac{\left(\sum_{i=1}^a \sum_{j=1}^b r_{ij}^2 - \text{SEE}(\text{cluster})\right) / 2}{\text{SSE}(\text{cluster}) / ((a-1)(b-1)-2)}.$$

The hypothesis  $H_0$  is rejected for large values of  $T_{\text{Malik}}$ . Note that the result of the Malik test may depend on the method of clustering. In the **combinIT** package, clustering is done by *kmeans* function in **RcppArmadillo**. The *speed\_mode* parameter on the *kmeans* clustering was set as *static\_subset*. Note that the Malik test performs well when there are some cells that produce large negative or positive residuals due to significant interaction. The package **combinIT** provides the value of test statistic  $T_{\text{Malik}}$  and an exact p-value of the Malik test based on a Monte Carlo simulation procedure.

**KKM.** Let  $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$  denote the  $(i, j)$ -th cell mean and let  $\eta_{ij} = \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'}, i \neq i'; j \neq j'$ , denote pairwise interaction contrasts. Kharrati-Kopaei and Miller (2016) used the idea that if a significant interaction is present, its corresponding interaction contrast shall be significant. To describe this method, let  $\mathbf{Z}$  be an  $n \times 1$  vector with elements  $y_{ij} - y_{i'j} - y_{ij'} + y_{i'j'}$  where  $n = ab(a-1)(b-1)/4$  is the total number of estimated pairwise interaction contrasts. Kharrati-Kopaei and Miller (2016) proposed as a test statistic, where  $Z_k$  is the  $k$ -th element of  $\mathbf{Z}$ ,  $\text{PSE} = \text{median}\{|Z_k|; |Z_k| \leq 5S_0\}$  in which  $S_0 = \text{median}\{|Z_k|; k = 1, \dots, n\}/c_0$  where  $c_0$  is an unbiased constant obtained by a Monte Carlo simulation.  $H_0$  is rejected when  $T_{\text{KKM}}$  is large. This test procedure is powerful for detecting significant interactions when they are caused by some cells; i.e. some values of  $|Z_k|$  are large. The package **combinIT** provides the value of test statistic  $T_{\text{KKM}}$  and the exact p-value of the KKM test based on a Monte Carlo simulation procedure.

#### Four methods for combining interaction tests

Since there is no single best interaction test in an unreplicated two-way ANOVA model, it is reasonable to combine multiple interaction tests to provide researchers with a testing approach that leverages many existing methods to detect different patterns of non-additivity. In this section, we review four combination methods that were proposed by Shenavari and Kharrati-Kopaei (2018). Throughout the section, we assume that  $K \geq 2$  tests, whose test statistics are dependent and whose distribution functions are continuous, have been performed for a null hypothesis  $H_0$ . Therefore, there are  $K$  dependent p-values  $P_1, \dots, P_K$  corresponding to the interaction tests. We are going to combine these p-values into a single p-value using one of the following four test procedures for testing  $H_0$ .

**Bon.** Following the Bonferroni inequality,  $H_0$  can be rejected at level  $\alpha$  when  $Kp_{\min} < \alpha$  where  $p_{\min}$  is the observed value of  $P_{\min} = \min_{1 \leq i \leq K} \{P_i\}$ . Note that the Bonferroni method leads to a conservative method of combination; i.e. its Type I error rate is always less than  $\alpha$ ;  $\Pr_{H_0}\{KP_{\min} < \alpha\} \leq \alpha$ .

**Sidak.** The distribution function of  $P_{\min}$  under  $H_0$  would be a beta distribution with parameters 1 and  $K$  if  $P_i, i = 1, \dots, K$ , were independent; i.e.  $\Pr_{H_0}\{P_{\min} \leq x\} = 1 - (1-x)^K$ . When  $P_i$ s are dependent, it can be shown that there is an  $x^* \in (0, 1)$  such that  $\Pr_{H_0}\{P_{\min} \leq x\} \leq 1 - (1-x)^K$  for all  $x < x^*$ ; see Sadooghi-Alvandi and Kharrati-Kopaei (2015). Therefore,  $H_0$  would be rejected at level  $\alpha$  when  $1 - (1 - p_{\min})^K < \alpha$ .

**JPE.** The distribution of  $P_{\min}$  under  $H_0$  can be approximated by the Jacobi polynomial expansion. In this expansion, the distribution of  $P_{\min}$  is approximated by the weighted sum of several Beta distribution functions with parameters  $p = 1$  and  $q = 1/E_1 - 1$  where  $E_1$  denotes the first moment of  $P_{\min}$ ; see Shenavari and Kharrati-Kopaei (2018). Up to the second term in the Jacobi polynomial expansion, the distribution of  $P_{\min}$  is approximated by a Beta distribution with parameters 1 and  $q$ . In practice, the parameter  $q$  can be estimated as  $\hat{q} = \min\{\max\{1, 1/P_{\min} - 1\}, K - 1\}$ ; see Shenavari and Kharrati-Kopaei (2018). By this approximation, the null hypothesis  $H_0$  is rejected at level  $\alpha$  when  $1 - (1 - p_{\min})^{\hat{q}} < \alpha$ .

**GC.** Let  $Z_i = \Phi^{-1}(P_i), i = 1, \dots, K$ , where  $\Phi^{-1}(\cdot)$  is the quantile function of the standard normal distribution. Although it is well known that  $Z_i$ s are marginally distributed as the standard normal distribution under the null hypothesis  $H_0$ , the joint distribution of the  $Z_i$  is unknown in practice. One can use the Gaussian copula (GC) family to model the joint distribution of  $Z_i$ s. Therefore, one can assume that the random vector  $(Z_1, \dots, Z_K)$  has a  $K$ -variate normal distribution under the null hypothesis  $H_0$  with zero mean vector and variance-covariance matrix  $\Sigma$ . In this case, the null hypothesis  $H_0$  is rejected at level  $\alpha$  when

$$\Pr_{H_0}\{P_{\min} \leq p_{\min}\} = 1 - \Pr_{H_0}\{Z_1 > \Phi^{-1}(p_{\min}), \dots, Z_K > \Phi^{-1}(p_{\min})\} < \alpha.$$

This probability is calculated by using **mvtnorm** in the package **combinIT**. Note that  $\Sigma$  can be estimated by  $\hat{\Sigma} = (1 - \hat{\rho}) \mathbf{I}_K + \hat{\rho} \mathbf{1}_K \mathbf{1}'_K$  where  $\mathbf{I}_K$  is a  $K \times K$  identity matrix,  $\mathbf{1}_K$  is a  $K$ -dimensional vector of ones,

and  $\hat{\rho} = \max\{-1/(K-1), 1 - \sum_{i=1}^K (Z_i - \bar{Z})^2 / (K-1)\}$ ; see Shenavari and Kharrati-Kopaei (2018).

**Note 1.** If the null hypothesis  $H_0$  is rejected, the combined tests can suggest a pattern of non-additivity by examining the interaction test for which  $P_{\min} = \min_{1 \leq i \leq K} \{P_i\}$  occurs.

**Note 2.** The inequality  $Kp_{\min} \leq 1 - (1 - p_{\min})^K \leq 1 - (1 - p_{\min})^{\hat{q}}$  guarantees that the JPE method provides an interaction test that is never less powerful and is sometimes much more powerful than the Bon and Sidak counterparts.

### 3 Technical details about the **combinIT** package

The **combinIT** package stands for "Combined Interaction Test". This package allows users to perform the six NFBF interaction tests individually and to perform a combined test of the six NFBF interaction tests. It also has a function to produce an interaction plot and four data sets. The available functions in the **combinIT** package are described in Table 1. This package produces two S3-class objects: (i) **ITtest**, produced by functions that perform a single test and (ii) **combtest**, produced by the function that performs the combined test. These returned objects are lists with the following components:

- (i) **ITtest** class: For the return object of the functions `Malik_test`, `Franck_test`, `Boik_test`, `KKSA_test`, `KKM_test`, and `Piepho_test`. An object of this class consists of a list of six components, including:
  - `pvalue_exact`: The calculated exact Monte Carlo p-value. For `Boik_test`, this is an exact Monte Carlo p-value when  $p > 2$  and it is an exact p-value for  $p = 2$ .
  - `pvalue_appro`: For `Piepho_test` this is an asymptotic p-value. This component is the Bonferroni-adjusted p-value for `Franck.test` and `KKSA_test`. This component is not available for `Malik_test` and `KKM_test`.
  - `nsim`: The number of Monte Carlo samples used to calculate the exact Monte Carlo p-values.
  - `statistic`: The value of the test statistic.
  - `data_name`: The name of the input data set.
  - `test`: The name of the test.
- (ii) **combtest** class: For the return object of the function `CI_test`. An object of this class consists of a list of 19 components, including:
  - `nsim`: The number of Monte Carlo samples used to calculate the exact Monte Carlo p-values.
  - `Piepho_pvalue`: The p-value of Piepho's (1994) test.
  - `Piepho_Stat`: The value of Piepho's (1994) test statistic.
  - `Boik_pvalue`: The p-value of Boik test.
  - `Boik_Stat`: The value of Boik's (1993) test statistic.
  - `Malik_pvalue`: The p-value of Malik, et al.'s (2016) test.
  - `Malik_Stat`: The value of Malik, et al.'s (2016) test statistic.
  - `KKM_pvalue`: The p-value of Kharrati-Kopaei and Miller's (2016) test.
  - `KKM_Stat`: The value of Kharrati-Kopaei and Miller's (2016) test statistic.
  - `KKSA_pvalue`: The p-value of Kharrati-Kopaei and Sadooghi-Alvandi's (2007) test.
  - `KKSA_Stat`: The value of Kharrati-Kopaei and Sadooghi-Alvandi's (2007) test statistic.
  - `Franck_pvalue`: The p-value of Franck, et al.'s (2013) test.
  - `Franck_Stat`: The value of Franck, et al.'s (2013) test statistic.
  - `Bonferroni`: The combined p-value by using the Bonferroni method.
  - `Sidak`: The combined p-value by using the Sidak method.
  - `Jacobi`: The combined p-value by using the Jacobi method.
  - `GC`: The combined p-value by using the Gaussian copula.
  - `data_name`: The name of the input data set.
  - `test`: The name of the test.

The `print` method is developed for a brief displaying of these objects. One of the key features of this package is to make it possible for users to obtain accurate results for various tests within a reasonable time. This feature was achieved by assigning the time-consuming parts of these tests to several C++ functions that are called from R using **Rcpp**.

Function	Descriptions	Inputs	Main outputs
Boik_test	Performs the Boik test	The data matrix, the number of Monte Carlo samples for calculating the exact p-value.	The value of test statistic, an asymptotic p-value, the exact Monte Carlo p-value.
Piepho_test	Performs the Piepho test	The data matrix, the number of Monte Carlo samples for calculating an exact p-value.	The value of test statistic, an asymptotic p-value, the exact Monte Carlo p-value.
KKSA_test	Performs the KKSA test	The data matrix, the number of Monte Carlo samples for calculating the exact p-value, logical Elapsed_time for printing the progress.	The value of test statistic, an adjusted p-value, the exact Monte Carlo p-value.
Franck_test	Performs the Franck test	The data matrix, the number of Monte Carlo samples for calculating the exact p-value, logical Elapsed_time for printing the progress.	The value of test statistic, an adjusted p-value, the exact Monte Carlo p-value.
Malik_test	Performs the Malik test	The data matrix, the number of Monte Carlo samples for calculating the exact p-value, logical Elapsed_time for printing the progress.	The value of test statistic and the exact Monte Carlo p-value.
KKM_test	Performs the KKM test	The data matrix, the number of Monte Carlo samples for calculating the exact p-value and the unbiasing constant.	The value of test statistic and the exact Monte Carlo p-value.
CI_test	Performs the six NFBF tests and their combined tests	The data matrix, the number of Monte Carlo samples for calculating the exact p-value and the unbiasing constant, logical Elapsed_time for printing the progress.	The value of six test statistics, the exact Monte Carlo p-values of tests, and the result of four combined tests.
interaction_plot	Plots the interaction plot	The matrix of data.	Interaction plot.

**Table 1:** A summary of the functions in the **combinIT** package .

## 4 Using the package combinIT

Here, we illustrate the usage of the **combinIT** package. We also compare the new package with **hiddenf** in terms of execution speed. For reference, the functions of **combinIT** have been shown in Table 1. This table also presents some information about the inputs and outputs of functions.

We consider the copy number variation (CNV) data set that contains CNV values for normal and tumor tissue samples among six dogs. In this data set, the value of CNV was measured as a signal intensity obtained from a comparative genomic hybridization (CGH) array, with higher signals corresponding to higher copy numbers; see [Franck et al. \(2013\)](#) and [Osborne et al. \(2016\)](#). The following code loads the data and plots the interaction plot. It can be seen from [Figure 1](#) that a hidden structure might exist (rows 4 and 5 are parallel and they interact with rows 1, 2, 3, and 6).

```
> library(combinIT)
> data(CNV)
> interaction_plot(CNV)
```

Now, we use `CI_test` to combine the result of the six interaction tests to see if there exists any significant interaction. We note that the exact Monte Carlo p-values are obtained based on 10000 Monte Carlo samples and hence the absolute error of Monte Carlo approximation is at most 0.0082 with the 0.95 confidence coefficient (following the Central Limit Theorem). In addition, the p-value of the Boik test will be 1 since  $p = 1$ . We also set  $\alpha = 0.05$  (as default to test at the 5% level) and `report=TRUE` (to provide a report on the test), and `opvalue=0.6187` as the p-value of the Tukey's test (that is, we are going to combine the p-value of the Tukey's test in addition to the six interaction tests).

```
> CI_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE, ncol = 10000,
opvalue = 0.6187, Elapsed_time = FALSE)
Test: Combined interaction Test
Data: CNV
Piepho Test: Statistic = 1.73949 , Pvalue = 0.8743
```

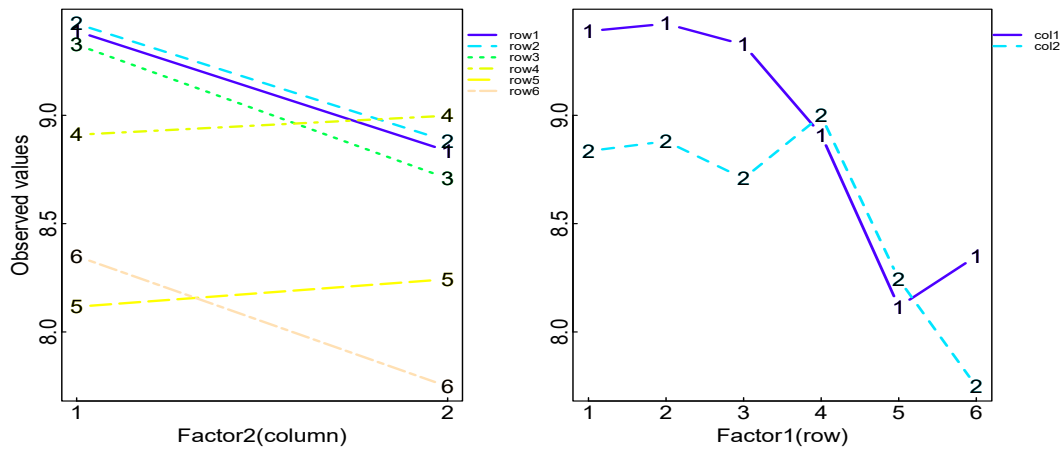


Figure 1: The interaction plot of the CNV data.

```

Boik Test: Statistic = 1 , Pvalue = 1
Malik Test: Statistic = 526.66767 , Pvalue = 0.0289
KKM Test: Statistic = 1.18173 , Pvalue = 0.9994
KKSA Test: Statistic = 0.0112 , Pvalue = 0.2092
Franck Test: Statistic = 526.66767 , Pvalue = 7e-04
Bonferroni method: Pvalue = 0.0049
Sidak method: Pvalue = 0.00489
Jacobi method: Pvalue = 0.00419
Gaussian copula: Pvalue = 0.00491
    
```

```

-----
A report on the combined interaction test:
A significant hidden structure exists at the 5% level. The first
group includes rows: 4, 5. The second group includes rows: 1, 2, 3,
6. The estimated critical value of the Franck_test at the 5% level
with 10000 Monte Carlo samples is 57.327.
    
```

The CI\_test function has detected a significant interaction at the 5% level and the detected interaction type is the hidden structure as we expect. All combined tests report that there is a significant interaction. It is worth mentioning that a significant interaction would not be detected if only Boik or Piepho or KKM or KKSA test was used. However, the combined tests provide researchers with much more opportunity to detect a significant interaction. One can use the Franck\_test to analyze the hidden structure in the CNV data set in more detail. The argument plot was set as TRUE to plot the interaction plot in the Franck\_test function. This interaction plot in Figure 2 shows that rows 4 and 5 are parallel (plotted in blue and solid line) and that they interact with rows 1, 2, 3, and 6 (plotted in red and broken line).

```

> Franck_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE, plot = TRUE,
Elapsed_time = FALSE)
Test: Franck Test
Data: CNV
Statistic = 526.668
Exact Monte Carlo P-value = 0
Approximate P-value = 0.001
Nsim = 10000
    
```

```

-----
A report on the test:
A significant hidden structure exists at the 5% level. The first
group includes rows: 4, 5. The second group includes rows: 1, 2, 3,
6. The estimated critical value of the Franck_test at the 5% level
with 10000 Monte Carlo samples is 56.7246.
    
```

Although the other five interaction tests can be used individually to test interaction, one should take care that a Type I error has not resulted from the multiple tests that have been performed. Code for each test is shown below.

```

> Boik_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE)
Test: Boik Test
    
```

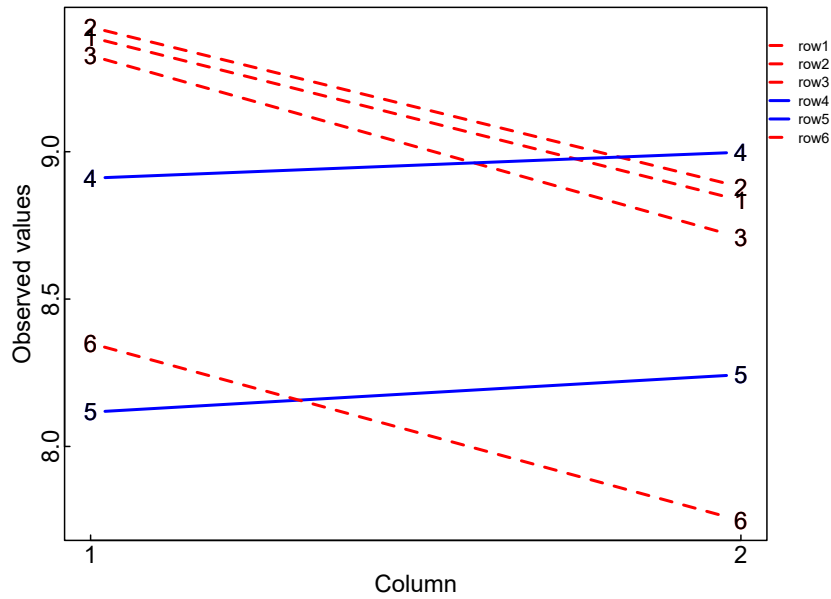


Figure 2: The interaction plot of the CNV data to see the hidden structure of interaction.

```

Data: CNV
Statistic = 1
Exact Monte Carlo P-value = 1
Approximate P-value = 1
Nsim = 10000
-----
A report on the test:
The Boik_test could not detect any significant interaction at the 5%
level. The exact critical value of the Boik_test is 1.
> Piepho_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE)
Test: Piepho Test
Data: CNV
Statistic = 1.739
Exact Monte Carlo P-value = 0.878
Approximate P-value = 0.884
Nsim = 10000
-----
A report on the test:
The Piepho_test could not detect any significant interaction at the
5% level. The estimated critical value of the Piepho_test at the 5%
level with 10000 Monte Carlo samples is 12.5169.
> KKSA_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE, plot = FALSE,
Elapsed_time = FALSE)
Test: KKSA Test
Data: CNV
Statistic = 0.011
Exact Monte Carlo P-value = 0.206
Approximate P-value = 0.28
Nsim = 10000
-----
A report on the test:
The KKSA_test could not detect any significant interaction at the 5%
level. The estimated critical value of the KKSA_test at the 5% level
with 10000 Monte Carlo samples is 0.0023.
> Malik_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE, Elapsed_time = FALSE)
Test: Malik Test
Data: CNV
Statistic = 526.668
Exact Monte Carlo P-value = 0.026
    
```



```

Approximate P-value = NULL
Nsim = 10000
-----
A report on the test:
There exists a significant interaction at the 5% level. The
significant interaction might be due to some outliers in residuals;
some cells produce large negative or positive residuals:
The cell with row=5 and column=1 produces a large negative residual
The cell with row=5 and column=2 produces a large positive residual
The estimated critical value of the Malik_test at the 5% level with
10000 Monte Carlo samples is 362.0247.
> KKM_test(CNV, nsim = 10000, alpha = 0.05, report = TRUE, nc0 = 10000)
Test: KKM Test
Data: CNV
Statistic = 1.182
Exact Monte Carlo P-value = 0.999
Approximate P-value = NULL
Nsim = 10000
-----
A report on the test:
The KKM_test could not detect any significant interaction at the 5%
level. The estimated critical value of the KKM_test at the 5% level
with 10000 Monte Carlo samples is 4.2359.

```

As we see, `Boik_test`, `Piepho_test`, `KKSA_test`, and `KKM_test` could not detect any significant interaction at the 5% level. However, `Malik_test` is significant and the significant interaction is due to the cells (5,1) and (5,2) that produce large negative and positive residuals, respectively. Note that the `KKSA_test` has an argument `plot` that if `plot` is `TRUE` an interaction plot will be plotted. Color and line type are used to display which levels of row factor are assigned to which sub-tables based on the minimum p-values among all possible configurations. The values of the Monte Carlo p-value and the adjusted p-value of the Franck test differ little (the presented values were rounded up to three decimal digits and the actual value of the Monte Carlo p-value is around 0.00069); on the other hand, they differ for the `KKSA_test`. This shows the importance of calculating the exact Monte Carlo p-values, while the package `hiddenf` provides only the adjusted p-values.

We now compare the execution speed of the `Malik_test` function in the `combinIT` package with the `MalikPvalue` function in the `hiddenf` package. To do this, we consider the `CNV` data again. We use the command `system.time` in R to record the elapsed time of execution in seconds for comparing the speed of execution. The elapsed time for 500 (the default of the `hiddenf` package), 1000, 5000, 10000 (the default of the `combinIT` package), 50000, and 100000 Monte Carlo samples are shown in Figure 3. All computations were done on a laptop with Windows 10 operating system, Intel(R) Core(TM) i5-4200U CPU, 1.60GHz 2.30 GHz processor, and 8G of RAM. The following shows the R code used and the output for 500 Monte Carlo samples.

```

> library(hiddenf)
> CNV.mtx<-HiddenF(CNV)
> system.time(
+ Malik_test(CNV,nsim=500,Elapsed_time = FALSE)
+ )
user system elapsed
0.15 0.13 0.22
> system.time(
+ MalikPvalue(CNV.mtx,N=500)
+ )
(Pvalue from Malik's test estimated with N=500 Monte Carlo datasets)
user system elapsed
4.05 0.00 4.08

```

It is seen that the run time of the `Malik_test` function is much shorter than that of the `MalikPvalue` function. Actually, the run speed of the `Malik_test` function is at least 18.5 times faster than the run speed of the `MalikPvalue` function.

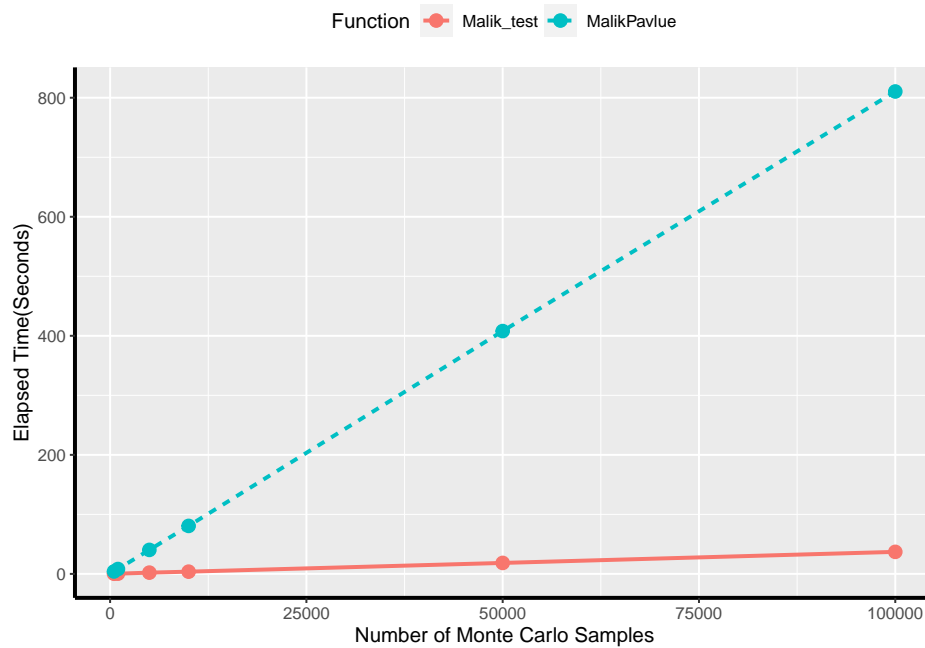


Figure 3: The elapsed time of executing MalikPvalue and Malik\_test functions for the CNV data.

## 5 Simulation study

In this section, we examine the performance of the Bon, Sidak, JPE, and GC methods in terms of controlling the Type I error rate using the **combinIT** and **hiddenf** packages. Note that [Shenavari and Kharrati-Kopaei \(2018\)](#) evaluated the performance of these methods for combining the p-values of the Boik, Piepho, KKSA, Franck, KKM, and Malik tests. Here, in addition to the six p-values, we consider the p-values of Tukey's test and Mandel's test ([Tukey, 1949](#); [Mandel, 1961](#)) as the two other tests that may apply. We used the following procedure to estimate the Type I error rate of the combination methods. For given  $a$  and  $b$ , an  $a \times b$  data table is generated from the standard normal distribution. We also took  $\sigma^2 = 1$  without loss of generality (since the tests are scale invariant). For the computed interaction effects" changes to "For the generated data table, the tests are done and their p-values are recorded. For each method of combination, these p-values are combined and the result of the combined test's rejection of the null hypothesis of no interaction is recorded. This process is repeated  $N1$  times and the fraction of times that the combined test rejects the null hypothesis is calculated as an estimate of the Type I error rate of the combining method.

The results of simulation for  $\alpha = 0.10, 0.05, 0.01$ ,  $N1 = 10,000$ , and different values of  $a$  and  $b$  are shown in Table 2. The exact Monte Carlo p-values of tests were calculated based on  $N2 = 100,000$  Monte Carlo samples. In this table, the estimated absolute errors (EAE) in estimating the Type I error rate with confidence coefficient 0.95 are shown in parentheses. The value of EAE with confidence coefficient 0.95 is  $1.96\sqrt{\hat{\alpha}(1-\hat{\alpha})/N1}$  where  $\hat{\alpha}$  denotes the estimated Type I error rate.

It is seen from Table 2 that the combining methods control the Type I error rate. This shows that these methods can be used for combining the p-values of the other tests successfully that may be introduced in the future. Among these four combination methods, the JPE method performs better than the others since the values of its estimated Type I error are fairly close to the nominal Type I error. We did not examine the performance of methods in terms of detecting a significant interaction power, because the considered methods do not have the same estimated Type I error rate (The JPE method would have the highest power, because its estimated Type I error rate is higher than the other methods; see also Note 2).

## 6 Summary

We reviewed six recommended interaction tests in the context of unreplicated two-way ANOVA models and four combination methods for combining dependent interaction tests. Then, we introduced the **combinIT** package and demonstrated that the advantages of this package over the two existing packages are: (i) it reports test statistics, estimated critical value, approximate p-values, and exact

$\alpha$	$a \times b$	Methods			
		Bon	Sidak	JPE	GC
0.10	$4 \times 3$	0.0681(0.0049)	0.0716(0.0051)	0.0822(0.0054)	0.0770(0.0052)
	$5 \times 3$	0.0738(0.0051)	0.0771(0.0052)	0.0876(0.0055)	0.0784(0.0054)
	$5 \times 5$	0.0752(0.0052)	0.0792(0.0053)	0.0879(0.0055)	0.0799(0.0053)
	$6 \times 6$	0.0723(0.0051)	0.0749(0.0052)	0.0849(0.0055)	0.0764(0.0052)
0.05	$4 \times 3$	0.0372(0.0037)	0.0377(0.0037)	0.0423(0.0039)	0.0388(0.0038)
	$5 \times 3$	0.0384(0.0038)	0.0394(0.0038)	0.0441(0.0040)	0.0396(0.0038)
	$5 \times 5$	0.0399(0.0038)	0.0411(0.0038)	0.0463(0.0041)	0.0406(0.0039)
	$6 \times 6$	0.0395(0.0038)	0.397(0.0038)	0.0448(0.0041)	0.0402(0.0038)
0.01	$4 \times 3$	0.0090(0.0019)	0.0090(0.0019)	0.0103(0.0020)	0.0091(0.0019)
	$5 \times 3$	0.0078(0.0017)	0.0078(0.0017)	0.0088(0.0018)	0.0079(0.0017)
	$5 \times 5$	0.0090(0.0019)	0.0092(0.0019)	0.0101(0.0020)	0.0092(0.0019)
	$6 \times 6$	0.0087(0.0018)	0.0087(0.0018)	0.0096(0.0019)	0.0086(0.0018)

**Table 2:** The estimated Type I error rates of the combined tests and their EAE (the values in parentheses).

Monte Carlo p-values of six recommended NFBF interaction tests, (ii) it provides the results of the four combined interaction tests in addition to some descriptions of detected significant interaction patterns, and (iii) its execution is fast enough to make it feasible to calculate the exact Monte Carlo p-values. Using the **combinIT** package, we evaluated the performance of the four combined tests in terms of controlling the Type I error rate. Simulation results show that the combined tests control the Type I error rate, and therefore can be used as an interaction test that leverages existing methods to detect different patterns of non-additivity.

We note that the **combinIT** package can handle data sets as large as  $15 \times 15$ , although performing the KKSA and Franck tests may be time-consuming when the size of the data set is larger. We will focus on solving this problem in the next version of the package.

## 7 Acknowledgment

The first author would like to acknowledge the Research Council of Shiraz University.

## Bibliography

- A. Alin and S. Kurt. Testing non-additivity (interaction) in two-way anova tables with no replication. *Statistical Methods in Medical Research*, 15(1):63–85, 2006. URL <https://doi.org/10.1191/0962280206sm426oa>. PMID: 16477949. [p284]
- R. J. Boik. Testing additivity in two-way classifications with no replications: the locally best invariant test. *Journal of Applied Statistics*, 20(1):41–55, 1993a. URL <https://doi.org/10.1080/02664769300000004>. [p284, 285, 286]
- R. J. Boik. A comparison of three invariant tests of additivity in two-way classifications with no replications. *Computational Statistics and Data Analysis*, 15:411–424, 1993b. URL [https://doi.org/10.1016/0167-9473\(93\)90173-Q](https://doi.org/10.1016/0167-9473(93)90173-Q). [p285, 286]
- L. C. A. Corsten and A. C. v. Eijnsbergen. Multiplicative effects in two-way analysis of variance. *Statistica Neerlandica*, 26(3):61–68, 1972. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9574.1972.tb00173.x>. [p284]
- L. C. A. Corsten and A. C. v. Eijnsbergen. Addendum to “multiplicative effects in two-way analysis of variance”. *Statistica Neerlandica*, 28(1):51–51, 1974. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9574.1974.tb00231.x>. [p284]

- D. Eddebuettel. *Seamless R and C++ Integration with Rcpp*. Springer, New York, 2013. URL <https://doi.org/10.1007/978-1-4614-6868-4>. ISBN 978-1-4614-6867-7. [p285]
- D. Eddebuettel and J. J. Balamuta. Extending extitR with extitC++: A Brief Introduction to extitR-cpp. *The American Statistician*, 72(1):28–36, 2018. URL <https://doi.org/10.1080/00031305.2017.1375990>. [p285]
- D. Eddebuettel and R. François. Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18, 2011. URL <https://doi.org/10.18637/jss.v040.i08>. [p285]
- C. T. Franck and J. A. Osborne. Exploring interaction effects in two-factor studies using the hiddenf package in r. *The R Journal*, 8(1):159–172, 2016. URL <https://doi.org/10.32614/RJ-2016-011>. [p284, 285]
- C. T. Franck, D. M. Nielsen, and J. A. Osborne. A method for detecting hidden additivity in two-factor unreplicated experiments. *Computational Statistics and Data Analysis*, 67:95–104, Nov 2013. ISSN 0167-9473. URL <http://dx.doi.org/10.1016/j.csda.2013.05.002>. [p284, 285, 286, 289]
- D. E. Johnson and F. A. Graybill. An analysis of a two-way model with interaction and no replication. *Journal of the American Statistical Association*, 67(340):862–868, 1972. ISSN 01621459. URL <http://www.jstor.org/stable/2284651>. [p284, 285]
- M. Kharrati-Kopaei and A. Miller. A method for testing interaction in unreplicated two-way tables: using all pairwise interaction contrasts. *Journal of Statistical Computation and Simulation*, 86(6):1203–1215, 2016. URL <https://doi.org/10.1080/00949655.2015.1057821>. [p284, 285, 287]
- M. Kharrati-Kopaei and S. M. Sadooghi-Alvandi. A new method for testing interaction in unreplicated two-way analysis of variance. *Communications in Statistics - Theory and Methods*, 36:2787 – 2803, 2007. URL <https://doi.org/10.1080/03610920701386851>. [p284, 285, 286]
- W. A. Malik, J. Möhring, and H. P. Piepho. A clustering-based test for nonadditivity in an unreplicated two-way layout. *Communications in Statistics - Simulation and Computation*, 45(2):660–670, 2016. URL <https://doi.org/10.1080/03610918.2013.870196>. [p284, 285, 286]
- J. Mandel. Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association*, 56(296):878–888, 1961. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1961.10482132>. [p284, 285, 293]
- J. Mandel. A new analysis of variance model for non-additive data. *Technometrics*, 13(1):1–18, 1971. URL <https://www.tandfonline.com/doi/abs/10.1080/00401706.1971.10488751>. [p284, 285]
- G. A. Milliken and D. E. Johnson. *Analysis of Messy Data, Volume II: Nonreplicated Experiments*. Chapman and Hall/CRC, 1st ed edition, 1989. URL <https://doi.org/10.1201/9781315172194>. [p284]
- G. A. Milliken and D. M. Rasmuson. A heuristic technique for testing for the presence of interaction in nonreplicated factorial experiments. *Australian and New Zealand Journal of Statistics*, 19:32–38, 1977. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-842X.1994.tb00889.x>. [p284, 285, 286]
- D. C. Montgomery. *Design and Analysis of Experiments*. John Wiley and Sons, Inc., Hoboken, NJ, USA, 2017. ISBN 9781119113478. [p284]
- J. A. Osborne, C. T. Franck, and B. Choi. *hiddenf: The All-Configurations, Maximum-Interaction F-Test for Hidden Additivity*, 2016. URL <https://CRAN.R-project.org/package=hiddenf>. R package version 2.0. [p285, 289]
- H.-P. Piepho. On tests for interaction in a nonreplicated two-way layout. *Australian Journal of Statistics*, 36(3):363–369, 1994. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-842X.1994.tb00889.x>. [p284, 285, 286]
- S. M. Sadooghi-Alvandi and M. Kharrati-Kopaei. Testing three-factor interaction in unreplicated three-way layouts. *Communications in Statistics - Theory and Methods*, 44(3):587–606, 2015. URL <https://doi.org/10.1080/03610926.2012.748915>. [p287]
- Z. Shenavari and M. Kharrati-Kopaei. A method for testing additivity in unreplicated two-way layouts based on combining multiple interaction tests. *International Statistical Review*, 86(3):469–487, 2018. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12262>. [p284, 287, 288, 293]

- Z. Shenavari, H. Haghbin, M. Kharrati-Kopaei, and S. M. Najibi. *combinIT: A Combined Interaction Test for Unreplicated Two-Way Tables*, 2022. URL <https://CRAN.R-project.org/package=combinIT>. R package version 2.0.0. [p285]
- P. Simecek and M. Simeckova. Modification of tukey's additivity test. *Journal of Statistical Planning and Inference*, 143:197–201, 2013. URL <https://doi.org/10.1016/j.jspi.2012.07.002>. [p284, 285]
- M. Simeckova, T. Rusch, and P. Simecek. *additivityTests: Additivity Tests in the Two Way Anova with Single Sub-class Numbers*, 2014. URL <https://CRAN.R-project.org/package=additivityTests>. R package version 1.1-4. [p285]
- J. W. Tukey. One degree of freedom for non-additivity. *Biometrics*, 5(3):232–242, 1949. ISSN 0006341X, 15410420. URL <http://www.jstor.org/stable/3001938>. [p284, 285, 293]
- F. Tusell. Testing for interaction in two-way anova tables with no replication. *Computational Statistics and Data Analysis*, 10(1):29–45, 1990. ISSN 0167-9473. URL <https://www.sciencedirect.com/science/article/pii/016794739090101M>. [p284, 285]

Mahmood Kharrati-Kopaei  
Department of Statistics, College of Science, Shiraz University  
Shiraz  
Iran  
(ORCID 0000-0001-5555-253X)  
[mkharati@shirazu.ac.ir](mailto:mkharati@shirazu.ac.ir)

Zahra Shenavari  
Department of Mathematics, Faculty of Sciences, Shiraz Branch, Islamic Azad University  
Shiraz  
Iran  
(ORCID 0000-0001-8347-0590)  
[zshenavari@yahoo.com](mailto:zshenavari@yahoo.com)

Hossein Haghbin  
Department of Statistics, Faculty of Intelligent Systems Engineering and Data Science, Persian Gulf University  
Boushehr  
Iran  
(ORCID 0000-0001-8416-2354)  
<https://github.com/haghbinh>  
[haghbin@pgu.ac.ir](mailto:haghbin@pgu.ac.ir)