Journal of Statistical Modelling: Theory and Applications
Vol. 6, No. 2, 2025, pp. 153-171
Yazd University Press 2025

Research Paper

Tests for one-way analysis of covariance with heteroscedasticity

Ali Akbar Jafari*,[1], Soltan Mohammad Sadooghi-Alvandi[2]
[1]Department of Statistics, Yazd University, Yazd, Iran
[2]Department of Statistics, Shiraz University, Shiraz, Iran

Received: July 23, 2025/ Revised: February 09, 2025/ Accepted: August 28, 2026

Abstract: The one-way analysis of covariance model is used to evaluate the equality between multiple treatments in the presence of a covariate. Presenting a test that controls the size is a major consideration for this model especially when the variances are unequal, since the actual size of the test depends on the values of the variances. The traditional F test in the analysis of covariance problem can produce unreliable outcomes, and there is no simple manipulation test that satisfactorily controls the size. The calculations for the generalized F test and parametric bootstrap test are complicated and must be performed using Monte Carlo techniques. In this study, we first suggest a generalized test for the analysis of covariance problem and then construct three approximate tests that do not have the problems of earlier tests. The performance of the generalized test and the three simple tests is shown using extensive simulations. When the variances are very unequal, the actual sizes of these tests can be close to the nominal size and rarely exceed the nominal level. Three data sets illustrate the application of the methods.

Keywords: Analysis of covariance; Behrens-Fisher problem; Generalized test; Heteroscedasticity.
Mathematics Subject Classification (2010): 60E05, 62F12.

1 Introduction

In an analysis of variance used to compare several populations, additional information about concomitant variables can reduce experimental errors. Therefore, the experimenter can control these variables and thus obtain a better analysis of the response

*Corresponding author: aajafari@yazd.ac.ir

variable by the inclusion of one or more covariates. In this situation, the analysis of covariance (ANCOVA) model is used for modeling the variables, employing a technique that combines regression and analysis of variance (Rutherford, 2011). The ANCOVA model has been utilized in various scientific research fields, such as educational research (Keselman et al., 1998), agriculture (Milliken and Johnson, 2001), psychology experiments (Rutherford, 2011), and clinical trials (European Medicines Agency, 2015). Barrett (2011) presented a review of computations of the one-way fixed effects ANCOVA model, the assumptions of the test, and techniques to illustrate those using generalized linear models. Milliken and Johnson (2001) provided good descriptions of different types of covariance analysis models and their applications. Khammar et al. (2020) presented some notes on using the ANCOVA methodology and the appropriate way of reporting the results in medical research.

In this paper, we consider the ANCOVA problem of comparing several treatments when there are one or more covariates and the variances are not considered to be equal. This model to compare k treatments with a single covariate can be presented as

$$Y_{ij} = \mu_i + \beta x_{ij} + \varepsilon_{ij}, \quad i = 1, \ldots, k, \quad j = 1, \ldots, n_i, \tag{1}$$

where Y_{ij} and x_{ij} denote the response and covariate values, respectively, μ_i and β are unknown parameters, and ε_{ij} are independent error terms with $\varepsilon_{ij} \sim N\left(0, \sigma_i^2\right)$. The main hypothesis of interest is

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k. \tag{2}$$

The traditional F test (Arnold, 1981) can be used to test the hypothesis in (2), if the variances are considered to be equal. However, this test can produce misleading results if the variances are not equal. Also, the method tends to be liberal or conservative, depending on the variances, sample sizes, and the shape of the distribution (Sadooghi-Alvandi and Jafari, 2013).

The ANCOVA testing problem is a Behrens-Fisher type problem if the variances are unequal; that is, no test has size for all values of the nuisance parameters σ_i^2. Since the actual size of the test (i.e., the probability of Type I error) depends on the values of the nuisance parameters σ_i^2, controlling the test size is an essential aspect in solving these types of problems. Ananda (1998) presented a generalized F test for the ANCOVA model with unequal variances based on the ideas of generalized test variable and generalized p-value (Tsui and Weerahandi, 1989). However, when sample sizes are small, particularly when the number of treatments is large, the actual size of this test can be significantly higher than the nominal level. Through simulation, Sadooghi-Alvandi and Jafari (2013) showed that the generalized F test cannot control the size as well as the PB test. The boldfaced cases in Tables 1-4 represent situations where the actual size of the PB test is greater than the nominal level. Additionally, the generalized F test and PB test require Monte Carlo methods for performing the necessary calculations, and there is no readily available simple test for the equality of treatment effects.

In this work, the generalized test variable and generalized p-value concepts are first used to develop a test. Then, using this generalized test as a basis, we suggest three approximate simple tests that do not have the problems that the generalized test has. These tests are similar to the Welch test (Welch, 1951) for the ANOVA problem and are of conventional form, and thus do not rely on the generalized test. Our proposed tests provide satisfactory solutions to the ANCOVA problem. For reasons of simplicity, we have focused just on one-way ANCOVA. However, our tests can be extended to the situation where there are several covariates.

The paper is organized as follows. We introduce our generalized test and develop its properties in Section 2. We use this test in Section 3 to suggest two simple F tests, as well as an adjusted chi-square test. We provide simulation results in Section 4 to illustrate how the generalized test and three simple tests perform. Additionally, three examples are used to illustrate these tests.

2 New generalized variable test

In this section, we introduce and explore the characteristics of a new generalized test variable. The proofs of some results are given in the Appendix.

2.1 Preliminaries

Suppose that $\bar{x}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$ and $\bar{Y}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$. The least square estimators of the main parameters of the ANCOVA model (1) are $\hat{\mu}_i = \bar{Y}_{i.} - \hat{\beta}\bar{x}_{i.}$, $i = 1, \ldots, k$, and

$$\hat{\beta} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_{i.}\right)\left(x_{ij} - \bar{x}_{i.}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{i.}\right)^2}.$$

Also, an unbiased estimator of σ_i^2 is

$$S_i^2 = \frac{1}{(n_i - 2)}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_{i.}\right)^2 - \frac{\left[\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_{i.}\right)\left(x_{ij} - \bar{x}_{i.}\right)\right]^2}{(n_i - 2)\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{i.}\right)^2}.$$

The statistics $S_1^2, \ldots, S_k^2$, and $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \ldots, \hat{\mu}_k)'$ are mutually independent, and $\hat{\boldsymbol{\mu}} \sim N_k(\boldsymbol{\mu}, \Sigma)$, while $S_i^2 \sim \frac{\sigma_i^2}{n_i-2}\chi^2_{(n_i-2)}$, $i = 1, \ldots, k$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)'$, $\bar{\boldsymbol{x}} = (\bar{x}_{1.}, \ldots, \bar{x}_{k.})'$, $\Sigma = \left[\mathrm{diag}\left(\frac{\sigma_i^2}{n_i}\right)\right] + \phi\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'$, $\phi = \sum_{i=1}^{k} l_i\sigma_i^2$, and $l_i = \sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{i.}\right)^2 / \left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_{i.}\right)^2\right]^2$.

The hypothesis H_0 in (2) may be expressed as $H_0 : C\boldsymbol{\mu} = \mathbf{0}$, where $C = [I_{k-1} : -\mathbf{1}_{k-1}]$, I_k is a $k \times k$ identity matrix, and $\mathbf{1}_m$ denotes an $1 \times m$ vector of 1's, $\mathbf{1}_m = (1, 1, \ldots, 1)'$. Under H_0, it can be shown that $(C\hat{\boldsymbol{\mu}})'(C\Sigma C')^{-1}(C\hat{\boldsymbol{\mu}})$ follows a chi-square distribution with $k-1$ degrees of freedom. Using this result, Sadooghi-Alvandi and Jafari (2013) proposed a test statistic with the following closed form

$$T_{\mathrm{SJ}} = \sum_{i=1}^{k}\frac{n_i}{S_i^2}\left(\hat{\mu}_i - \hat{\hat{\mu}}\right)^2 - \frac{\hat{\phi}\left[\sum_{i=1}^{k}\frac{n_i}{S_i^2}\left(\hat{\mu}_i - \hat{\hat{\mu}}\right)\left(\bar{x}_{i.} - \hat{\bar{x}}_{\omega}\right)\right]^2}{1 + \hat{\phi}\sum_{i=1}^{k}\frac{n_i}{S_i^2}\left(\bar{x}_{i.} - \hat{\bar{x}}_{\omega}\right)^2},$$

where $\hat{\bar{x}}_{\omega} = \frac{\sum_{i=1}^{k} n_i\bar{x}_{i.}/S_i^2}{\sum_{i=1}^{k}\bar{x}_{i.}/S_i^2}$, $\hat{\hat{\mu}} = \frac{\sum_{i=1}^{k} n_i\hat{\mu}_i/S_i^2}{\sum_{i=1}^{k} n_i/S_i^2}$, and $\hat{\phi} = \sum_{i=1}^{k} l_i S_i^2$. The chi-square distribution with $k-1$ degrees of freedom may approximate the null distribution of T_{SJ}

for large sample sizes. The parametric bootstrap approach was suggested by Sadooghi-Alvandi and Jafari (2013) as an alternative for this approximation because the latter performs poorly for small sample sizes. In Section 4, we demonstrate how to adjust the chi-square estimate for small sample sizes.

2.2 New generalized test

The natural estimator of Σ is $\hat{\Sigma} = \left[\text{diag}\left(\frac{S_i^2}{n_i}\right)\right] + \hat{\phi}\ \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'$. Let s_i^2, $\hat{\phi}_{\text{obs}}$, and $\hat{\Sigma}_{\text{obs}}$ denote the observed values of S_i^2, $\hat{\phi}$, and $\hat{\Sigma}$, respectively, that is, $\hat{\Sigma}_{\text{obs}} = \left[\text{diag}\left(\frac{s_i^2}{n_i}\right)\right] + \hat{\phi}_{\text{obs}}\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'$, and $\hat{\phi}_{\text{obs}} = \sum_{i=1}^{k} l_i s_i^2$. Based on the Cholesky decomposition, there is a unique lower triangular matrix L such that $LL' = C\hat{\Sigma}_{\text{obs}}C'$, because $C\hat{\Sigma}_{\text{obs}}C'$ is a positive definite matrix. Assuming $W = L^{-1}$, we have $W'W = \left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}$.

In the following, we follow a similar approach derived by Sadooghi-Alvandi et al. (2012) for the one-way ANOVA problem to test the equality of several normal means and propose a generalized test variable for the one-way ANCOVA problem. Let the variable $T_G = T_G\left(\hat{\mu}_i, S_i^2, s_i^2, \sigma_i^2\right)$ be defined as follows

$$\begin{aligned} T_G \quad &= \quad (C\hat{\boldsymbol{\mu}})'\, W'(WC\Sigma C'W')^{-\frac{1}{2}} WC \left\{ \left[\text{diag}\left(\frac{\sigma_i^2 s_i^2}{n_i S_i^2}\right)\right] + \left(\sum_{i=1}^{k} \frac{l_i \sigma_i^2 s_i^2}{S_i^2}\right) \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}' \right\} \\ &\quad \times C'W'\left(WC\Sigma C'W'\right)^{-\frac{1}{2}} W\left(C\hat{\boldsymbol{\mu}}\right). \end{aligned}$$

The observed value of T_G is then obtained as

$$t = (C\hat{\boldsymbol{\mu}}_{\text{obs}})'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}\left(C\hat{\boldsymbol{\mu}}_{\text{obs}}\right), \tag{3}$$

which is independent of the unknown parameters, where $\hat{\boldsymbol{\mu}}_{\text{obs}}$ is the observed value of $\hat{\boldsymbol{\mu}}$.

Theorem 2.1. *Under the null hypothesis, T_G given $s_1^2, \ldots, s_k^2$ is distributed as*

$$T^* = \boldsymbol{Z}'WC\left\{\left[\text{diag}\left(\frac{(n_i-2)s_i^2}{n_i U_i}\right)\right] + \left(\sum_{i=1}^{k}\frac{(n_i-2)\, l_i s_i^2}{U_i}\right)\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right\}C'W'\boldsymbol{Z},$$

where $\boldsymbol{Z} \sim N_{k-1}(\mathbf{0}, I_{k-1})$ and $U_i \sim \chi^2_{(n_i-2)}$, $i = 1, \ldots, k$, are mutually independent.

Proof. We may write

$$T_G = \boldsymbol{Z}'WC\left\{\left[\text{diag}\left(\frac{(n_i-2)s_i^2}{n_i U_i}\right)\right] + \left(\sum_{i=1}^{k}\frac{(n_i-2)\, l_i s_i^2}{U_i}\right)\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right\}C'W'\boldsymbol{Z},$$

where $\boldsymbol{Z} = (WC\Sigma C'\mathrm{W}')^{-\frac{1}{2}}WC\hat{\boldsymbol{\mu}}$ and $U_i = \frac{(n_i-2)S_i^2}{\sigma_i^2}$, $i = 1, \ldots, k$. Then, $U_1, \ldots, U_k$ are mutually independent, with $U_i \sim \chi^2_{(n_i-2)}$. Also, conditionally on s_i^2, the distribution of $\boldsymbol{Z}$ is multivariate normal, with mean zero (since $C\boldsymbol{\mu} = \mathbf{0}$ under H_0) and variance-covariance matrix $\text{Cov}\,(\boldsymbol{Z}) = (WC\Sigma C'W')^{-\frac{1}{2}}WC\Sigma C'W(WC\Sigma C'W')^{-\frac{1}{2}} = I_{k-1}$. Hence, $\boldsymbol{Z} \sim N_{k-1}(\mathbf{0}, I_{k-1})$, independently of $U_1, \ldots, U_k$. $\square$

It follows from Theorem 2.1 that the distribution of T_G does not depend on the nuisance parameters, σ_i^2, $i = 1, \dots, k$. As a result,

$$\mathrm{GPV} = P\left(T_G \geq t | H_0\right), \tag{4}$$

gives the generalized p-value for testing the hypothesis. It should be noted that if GPV $< \alpha$, then the null hypothesis is rejected at level α, yielding a traditional, fixed-level test. This test will be referred to as the GPV test.

Here, we provide an explicit form for the distribution of T_G based on the following results. The remainder of the paper uses the notation described below:

$$w_i = \left(\frac{n_i}{s_i^2}\right) / \sum_{j=1}^{k} \frac{n_j}{s_j^2}, \quad q_i = \sqrt{w_i}, \quad \bar{x}_w = \sum_i w_i \bar{x}_{i.}, \quad a_i = \frac{\sqrt{n_i}}{s_i}\left(\bar{x}_{i.} - \bar{x}_w\right),$$

$$A = \sum_{i=1}^{k} a_i^2, \quad K = \frac{\hat{\phi}_{\mathrm{obs}}}{1 + A\hat{\phi}_{\mathrm{obs}}} = \frac{\sum_i l_i s_i^2}{1 + \left(\sum_i a_i^2\right)\left(\sum_i l_i s_i^2\right)}. \tag{5}$$

Theorem 2.2. *Let $Z_i^* \sim N(0,1)$, $i = 1, \dots, k+1$, and $U_i \sim \chi^2_{(n_i - 2)}$, $i = 1, \dots, k$, be independent variables, and let $D_i = \frac{s_i}{\sqrt{n_i}} Z_i^* - \bar{x}_{i.} \sqrt{\hat{\phi}_{\mathrm{obs}}} Z_{k+1}^*$, $i = 1, \dots, k$. Then, under H_0, the generalized test variable T_G is distributed as T, defined by*

$$T = \sum_{i=1}^{k} \left\{ \frac{n_i (n_i - 2)}{U_i s_i^2} \left[\left(D_i - \bar{D}_w\right) - \frac{s_i}{\sqrt{n_i}} K H a_i \right]^2 \right\} + H^2 [1 - KA]^2 \sum_{i=1}^{k} \frac{(n_i - 2) l_i s_i^2}{U_i}, \tag{6}$$

where w_i, a_i, A, and K are as defined in (5)*, and $H = \sum_{i=1}^{k} \frac{\sqrt{n_i}}{s_i} a_i \left(D_i - \bar{D}_w\right)$, $\bar{D}_w = \sum_i w_i D_i$.*

The generalized p-value in (4) can be calculated using Monte Carlo techniques, as follows, based on Theorem 2.2:

Algorithm 2.3. *Given the data:*
Step 1: *Compute t as given by* (3)*.*
Step 2: *Compute the values of s_i^2, l_i, w_i, a_i $(i = 1, \dots, k)$, A, and K, as defined by* (5)*.*
Step 3: *Generate $Z_i^* \sim N(0,1)$, $i = 1, \dots, k+1$, and $U_i \sim \chi^2_{(n_i-2)}$, $i = 1, \dots, k$, independently.*
Step 4: *Compute the value of T, as given by* (6)*.*
Step 5: *Repeat Steps 3 and 4 a large number of times, M (say, $M = 100{,}000$), and let $M(t)$ be the number of times that $T \geq t$. The GPV is estimated as $M(t)/M$.*

2.3 Mean and variance

We derive explicit formulas for the first two moments of T_G under the null distribution in the following theorems. These are useful for proposing approximate tests.

Theorem 2.4. *Suppose that $n_i > 4$ for all i. Then, under H_0 and given s_i^2, we have*

$$E\left(T_G\right) = \sum_{i=1}^{k} \frac{n_i - 2}{n_i - 4} \left[1 - w_i - K a_i^2 + A\left(1 - AK\right) l_i s_i^2\right].$$

Remark 2.5. *It is easily verified that* $\sum_{i=1}^{k}\left[1-w_i-Ka_i^2+A(1-AK)l_is_i^2\right]=k-1$. *Therefore,* $E\left(T_G\right)\geq\frac{\max\{n_i\}-2}{\max\{n_i\}-4}\left(k-1\right)$, *and if* $n_i=n$, *then* $E\left(T_G\right)=\left(k-1\right)\left(n-2\right)/\left(n-4\right)$. *This shows why approximating the null distribution of* T_{SJ} *by the chi-square distribution, does not work well for small sample sizes.*

Theorem 2.6. *Suppose that* $n_i>6$ *for all* i. *Then, under* H_0 *and given* s_i^2, *we have*

$$\begin{aligned}
\mathrm{Var}\left(T_G\right) &= 2\sum_{i=1}^{k}\frac{\left(n_i-2\right)^2\left(1-2w_i-2Ka_i^2\right)}{\left(n_i-4\right)\left(n_i-6\right)}\\
&\quad+2\sum_{i=1}^{k}\frac{\left(n_i-2\right)^2}{\left(n_i-4\right)^2\left(n_i-6\right)}\Big\{2\left(w_i+Ka_i^2\right)^2+4(1-KA)^2l_is_i^2a_i^2\\
&\quad+2A^2(1-KA)^2l_i^2s_i^4+\left[1-w_i-Ka_i^2+A\left(1-KA\right)l_is_i^2\right]^2\Big\}\\
&\quad+2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}w_i\right)^2+4K\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}q_ia_i\right)^2+2K^2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}a_i^2\right)^2\\
&\quad+2A^2(1-KA)^2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}l_is_i^2\right)^2\\
&\quad+4(1-KA)^2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}l_is_i^2\right)\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}a_i^2\right).
\end{aligned}$$

3 Three approximate tests

In this section, we provide three simple tests that do not require the computations necessary for tests that employ the GPV technique or the PB approach. These tests are constructed by approximating the null distribution of the generalized test variable T_G. They would essentially provide approximations to the null distribution of T_{SJ}, leading to conventional tests based on T_{SJ}, because T_G has the same observed value as the statistic T_{SJ}.

We assume that $n_i>4$ for all i. This assumption is required to guarantee that the null distribution of T_G has a finite mean. It should be noted that this condition is acceptable since each treatment involves two unknown parameters, μ_i and σ_i^2, and thus at least five observations per treatment are needed. The (conditional) mean and variance of T_G under the null hypothesis, as stated in Theorems 2.4 and 2.6, are denoted as m_G and v_G in the following;

$$m_G=E\left(T_G|H_0\right),\quad v_G=\mathrm{Var}\left(T_G|H_0\right).$$

Let $\lambda_1,\ \ldots,\lambda_{k-1}$ ($\lambda_i>0$) denote the eigenvalues of

$$M=WC\left\{\left[\mathrm{diag}\left(\frac{(n_i-2)s_i^2}{n_iU_i}\right)\right]+\left(\sum_{i=1}^{k}\frac{\left(n_i-2\right)l_is_i^2}{U_i}\right)\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right\}C'W'.$$

Similar to Sadooghi-Alvandi et al. (2012), it can be shown that the null distribution of T_G has a representation of the form $T_G\sim\sum_{i=1}^{k-1}\lambda_iU_i^*$, where U_i^*'s are independent $\chi^2_{(1)}$ variables. This form suggests a chi-square approximation of the form $T_G \sim c^* \chi^2_{(k-1)}$. The value of c^* may then be obtained by equating the means, $m_G = E\left(c^* \chi^2_{(k-1)}\right)$, which gives $c^* = m_G/(k-1)$. The traditional test that results from this approximation is as follows:

The adjusted chi-square test. The null hypothesis is rejected at level α if

$$T_{\mathrm{SJ}} > \frac{m_G}{k-1} \chi^2_{(k-1)}(\alpha).$$

Regarding this test, recall from Section 2 that for large sample sizes, the null distribution of T_{SJ} is approximately $\chi^2_{(k-1)}$. This test adjusts for small sample sizes. For example, when the sample sizes are equal to 5, the adjustment factor is 3 (since $m_G = 3\,(k-1)$), and when the sample sizes are all equal to 6, the adjustment factor is 2 (since $m_G = 2\,(k-1)$). The results of our simulations, which are presented in Section 4, show that the adjusted chi-square test performs adequately in most cases.

To control the size, another approximation is obtained by using an F approximation of the form $T_G/(k-1) \sim F_{(k-1,r^*)}$. Again, the value of r^* is determined by equating the means, which gives $r^* = \frac{2m_G}{m_G-(k-1)}$. This approximation leads to the following F test:

The proposed F test I. Suppose $r^* = \frac{2m_G}{m_G-(k-1)}$. The null hypothesis is rejected at level α if

$$T_{\mathrm{SJ}} > (k-1)\, F_{(k-1,r^*)}\,(\alpha)\,.$$

If $n_i > 6$ for all i, the F approximation can be further improved by using an approximation of the form $T_G/(k-1) \sim cF_{(k-1,r)}$, with r and c obtained by equating both the means and the variances: $m_G = c\,(k-1)\,E\left(F_{(k-1,r)}\right)$ and $v_G = c^2(k-1)^2 \mathrm{Var}\left(F_{(k-1,r)}\right)$. This leads to the following F test:

The proposed F test II. Suppose $c = \frac{\left(m_G^2+v_G\right)m_G}{(k-3)m_G^2+2(k-1)v_G}$ and $r = 2\frac{(k-3)m_G^2+2(k-1)v_G}{(k-1)v_G-2m_G^2}$. The null hypothesis is rejected at level α if

$$T_{\mathrm{SJ}} > c\,(k-1)\, F_{(k-1,r)}\,(\alpha)\,.$$

4 Numerical studies

This section uses extensive simulations to first examine how well the tests control the size. Additionally, we demonstrate that if the comparison is based on the actual size rather than the nominal size, all five tests considered in this work have similar power properties. A more liberal test has greater power than a more conservative test in all other respects. In the end, three examples are presented.

4.1 Comparison of sizes

Here, we present the findings of a simulation study designed to evaluate the actual sizes of the PB test of Sadooghi-Alvandi and Jafari (2013), our GPV test, the adjusted chi-square test, and the two approximate F tests at the nominal level $\alpha = 0.05$. Note that the actual size of each test depends on the values of the nuisance parameters σ_i^2. The observations were generated from the ANCOVA model in (1) with $\mu_i = 10$ and $\beta = 1$, for different sample sizes n_i and values of variances σ_i^2. The values of the covariate were generated from $U(0, 10)$ and also from $N(5, 1)$. For each case, the p-values for the PB test and our GPV test were calculated using 10,000 repeats. The actual size of each test was then estimated by calculating the proportion of times (in $N = 100,000$ "runs") that the p-value was smaller than the nominal level $\alpha = 0.05$. We utilized the R programming language (R Core Team, 2025) to conduct our analyses.

The results of the simulations are presented in Tables 1–4. We can observe that

i. The actual size of the PB test is close to the nominal level $\alpha = 0.05$. However, in some situations, the actual size of this test exceeds the nominal level. We have highlighted these cases. In addition, the actual size of this test is larger than those of the other tests.

ii. The actual sizes of the suggested F tests and the GPV test were always less than the nominal level.

iii. Compared to F test I, the actual size of F test II (where applicable) was closer to the nominal level.

iv. The adjusted chi-square test has a larger actual size than the GPV test and the two proposed F tests.

v. The adjusted chi-square test rarely has a higher actual size than the nominal level. In most cases, it is close to or less than the nominal level.

vi. When the sample sizes of all groups are large, except for F test I, the actual sizes of all tests are close to the nominal level. The actual size of F test I is smaller than those of the other tests.

vii. The results are similar when the values of the covariate are generated from $U(0, 10)$ and $N(5, 1)$.

4.2 Power properties

In this section, we study the power properties of the PB test, the GPV test, the adjusted chi-square test, and our proposed F tests. Note that these five tests use the same statistic T_{SJ}, but with different approximations for its null distribution (which depends on the nuisance parameters σ_i^2). Therefore, a more liberal (conservative) test may <u>appear</u> to have higher (lower) power than a more conservative test. For example, the actual size of the proposed F test I is <u>lower</u> than the actual sizes of the other tests, and our extensive simulation (not reported in the paper) shows that the power of our proposed F test I is <u>lower</u> than the powers of the other tests. So, we compare the powers of the tests based on their <u>actual</u> sizes, rather than on the nominal sizes.

To illustrate, consider the case in Examples 4.1 and 4.2 with $k = 4$, $n_i = 7$, and $\beta = 1$. When testing at the nominal level $\alpha = 0.05$, the first column of Table 5 lists the actual sizes of the five tests. The tests should be applied at the nominal levels listed in the second column in order to achieve an actual size of 0.05; that is, when testing at these adjusted nominal levels, all tests have the same actual size of 0.05. Here, we consider two cases for the variances: i) $\sigma_1 = 1$, $\sigma_2 = 1.5$, $\sigma_3 = 2$, $\sigma_4 = 2.5$; ii) $\sigma_1 = 1$, $\sigma_2 = 100$, $\sigma_3 = 0.1$, $\sigma_4 = 0.01$. The powers of the tests (when testing at these "adjusted" nominal levels) are given in Table 5 for different values of $\mu_1, \ldots, \mu_4$. As expected, we found that the power properties of all five tests considered in this work are comparable.

Table 1: Actual sizes of the tests for $k = 4$ with generated covariates from $U(0, 10)$.

		$\sigma_1, \sigma_2, \sigma_3, \sigma_4$					
$n_1, \ldots, n_4$		1,1,1, 1	1,1.5,2,2.5	1,2,3,4	1,1,3,3	1,1,1,10	1,1,10,10
5,5,5,5	PB	0.0410	0.0451	0.0499	0.0485	0.0466	0.0534
	GPV	0.0184	0.0212	0.0228	0.0181	0.0258	0.0308
	Adj. chi-square	0.0231	0.0305	0.0304	0.0265	0.0331	0.0391
	Proposed F I	0.0158	0.0174	0.0236	0.0213	0.0196	0.0238
5,6,7,8	PB	0.0491	0.0451	0.0466	0.0465	0.0469	0.0484
	GPV	0.0258	0.0210	0.0242	0.0252	0.0282	0.0347
	Adj. chi-square	0.0367	0.0307	0.0343	0.0358	0.0390	0.0484
	Proposed F I	0.0172	0.0153	0.0215	0.0167	0.0171	0.0210
8,7,6,5	PB	0.0482	0.0547	0.0555	0.0558	0.0487	0.0499
	GPV	0.0210	0.0261	0.0333	0.0338	0.0370	0.0385
	Adj. chi-square	0.0307	0.0347	0.0434	0.0430	0.0429	0.0511
	Proposed F I	0.0175	0.0232	0.0290	0.0274	0.0282	0.0317
5,7,10,15	PB	0.0530	0.0484	0.0473	0.0492	0.0476	0.0501
	GPV	0.0354	0.0283	0.0271	0.0250	0.0262	0.0261
	Adj. chi-square	0.0425	0.0347	0.0328	0.0315	0.0305	0.0317
	Proposed F I	0.0212	0.0161	0.0178	0.0188	0.0180	0.0277
15,10,7,5	PB	0.0533	**0.0582**	**0.0581**	**0.0567**	0.0504	0.0541
	GPV	0.0277	0.0374	0.0434	0.0447	0.0420	0.0472
	Adj. chi-square	0.0358	0.0443	0.0503	0.0521	0.0456	0.0563
	Proposed F I	0.0188	0.0259	0.0320	0.0286	0.0288	0.0320
5,5,5,30	PB	**0.0559**	0.0525	**0.0555**	**0.0562**	0.0469	0.0531
	GPV	0.0258	0.0222	0.0216	0.0220	0.0159	0.0330
	Adj. chi-square	0.0328	0.0270	0.0275	0.0249	0.0186	0.0366
	Proposed F I	0.0179	0.0140	0.0175	0.0114	0.0105	0.0239
30,5,5,5	PB	**0.0564**	**0.0642**	**0.0626**	**0.0585**	**0.0516**	**0.0530**
	GPV	0.0252	0.0385	0.0386	0.0374	0.0251	0.0358
	Adj. chi-square	0.0317	0.0495	0.0479	0.0496	0.0332	0.0459
	Proposed F I	0.0171	0.0285	0.0326	0.0306	0.0232	0.0354
7,7,7,7	PB	0.0467	0.0485	0.0510	0.0505	0.0468	0.0507
	GPV	0.0282	0.0324	0.0304	0.0328	0.0361	0.0394
	Adj. chi-square	0.0401	0.0477	0.0435	0.0446	0.0494	0.0550
	Proposed F I	0.0203	0.0244	0.0268	0.0264	0.0295	0.0317
	Proposed F II	0.0238	0.0286	0.0265	0.0280	0.0320	0.0352
7,7,20,20	PB	0.0533	0.0505	0.0526	0.0500	0.0505	0.0504
	GPV	0.0385	0.0316	0.0378	0.0320	0.0320	0.0425
	Adj. chi-square	0.0491	0.0412	0.0471	0.0376	0.0419	0.0495
	Proposed F I	0.0273	0.0268	0.0225	0.0256	0.0260	0.0360
	Proposed F II	0.0301	0.0231	0.0293	0.0251	0.0256	0.0409
20,20,7,7	PB	0.0536	0.0544	0.0536	0.0529	0.0520	0.0510
	GPV	0.0382	0.0411	0.0426	0.0454	0.0414	0.0449
	Adj. chi-square	0.0514	0.0533	0.0578	0.0579	0.0544	0.0571
	Proposed F I	0.0321	0.0386	0.0390	0.0362	0.0319	0.0366
	Proposed F II	0.0284	0.0332	0.0345	0.0377	0.0330	0.0394
20,20	PB	0.0494	0.0496	0.0501	0.0500	0.0489	0.0492
20,20	GPV	0.0452	0.0486	0.0454	0.0455	0.0409	0.0447
	Adj. chi-square	0.0505	0.0536	0.0509	0.0514	0.0477	0.0508
	Proposed F I	0.0381	0.0365	0.0398	0.0373	0.0405	0.0425
	Proposed F II	0.0449	0.0488	0.0444	0.0452	0.0404	0.0449

Table 2: Actual sizes of the tests for $k = 6$ with generated covariates from $U(0, 10)$.

		$\sigma_1, \sigma_2, \ldots, \sigma_6$					
$n_1, \ldots, n_6$		1,1,1 1,1,1	1,1,1.5 2,2,2.5	1,2,3 1,2,3	1,1,1 3,3,3	1,1,1 1,1,10	1,1,1 10,10,10
5,5,5,5,5,5	PB	0.0394	0.0438	0.0472	0.0477	0.0441	0.0510
	GPV	0.0166	0.0158	0.0201	0.0208	0.0183	0.0291
	Adj. chi-square	0.0307	0.0304	0.0345	0.0342	0.0284	0.0439
	Proposed F I	0.0134	0.0137	0.0129	0.0162	0.0139	0.0231
5,6,7,8,9,10	PB	0.0501	0.0473	0.0485	0.0476	0.0506	0.0475
	GPV	0.0264	0.0245	0.0251	0.0215	0.0213	0.0254
	Adj. chi-square	0.0420	0.0414	0.0399	0.0333	0.0365	0.0402
	Proposed F I	0.0151	0.0134	0.0144	0.0107	0.0135	0.0165
10,9,8,7,6,5	PB	0.0511	**0.0547**	**0.0562**	**0.0533**	0.0477	0.0521
	GPV	0.0259	0.0310	0.0289	0.0353	0.0327	0.0435
	Adj. chi-square	0.0420	0.0485	0.0450	0.0549	0.0467	0.0598
	Proposed F I	0.0135	0.0243	0.0216	0.0205	0.0214	0.0289
5,5,7,7,10,15	PB	0.0520	0.0468	0.0512	0.0472	0.0481	0.0497
	GPV	0.0271	0.0196	0.0239	0.0214	0.0174	0.0246
	Adj. chi-square	0.0401	0.0312	0.0350	0.0344	0.0277	0.0375
	Proposed F I	0.0149	0.0101	0.0132	0.0122	0.0139	0.0127
10,10,10,5,5,5	PB	0.0524	**0.0570**	**0.0554**	**0.0605**	0.0508	0.0528
	GPV	0.0238	0.0324	0.0231	0.0406	0.0259	0.0395
	Adj. chi-square	0.0339	0.0445	0.0343	0.0546	0.0369	0.0573
	Proposed F I	0.0115	0.0219	0.0187	0.0281	0.0181	0.0318
5,5,5,5,5,30	PB	0.0509	0.0493	0.0509	0.0529	0.0434	0.0506
	GPV	0.0223	0.0206	0.0195	0.0221	0.0113	0.0317
	Adj. chi-square	0.0381	0.0330	0.0336	0.0347	0.0197	0.0455
	Proposed F I	0.0156	0.0124	0.0119	0.0156	0.0084	0.0228
30,5,5,5,5,5	PB	0.0506	**0.0557**	**0.0565**	**0.0565**	0.0473	0.0503
	GPV	0.0204	0.0276	0.0279	0.0305	0.0190	0.0339
	Adj. chi-square	0.0358	0.0456	0.0463	0.0478	0.0304	0.0502
	Proposed F I	0.0134	0.0221	0.0234	0.0238	0.0162	0.0252
7,7,7,7,7,7	PB	0.0483	0.0497	0.0490	0.0495	0.0475	0.0506
	GPV	0.0232	0.0301	0.0317	0.0294	0.0293	0.0344
	Adj. chi-square	0.0436	0.0497	0.0540	0.0495	0.0497	0.0571
	Proposed F I	0.0155	0.0202	0.0208	0.0212	0.0155	0.0240
	Proposed F II	0.0179	0.0243	0.0247	0.0220	0.0228	0.0283
7,7,10,12,15,15	PB	0.0511	0.0494	0.0508	0.0502	0.0499	0.0493
	GPV	0.0374	0.0329	0.0384	0.0350	0.0326	0.0356
	Adj. chi-square	0.0529	0.0495	0.0539	0.0485	0.0431	0.0485
	Proposed F I	0.0224	0.0222	0.0221	0.0190	0.0186	0.0213
	Proposed F II	0.0273	0.0252	0.0289	0.0255	0.0252	0.0289
20,20,20,7,7,7	PB	0.0541	0.0523	0.0531	0.0532	0.0524	0.0523
	GPV	0.0368	0.0450	0.0425	0.0451	0.0370	0.0460
	Adj. chi-square	0.0541	0.0592	0.0571	0.0588	0.0531	0.0653
	Proposed F I	0.0255	0.0305	0.0280	0.0314	0.0267	0.0341
	Proposed F II	0.0265	0.0317	0.0298	0.0333	0.0249	0.0354
20,20,20 20,20,20	PB	0.0504	0.0503	0.0502	0.0494	0.0496	0.0502
	GPV	0.0475	0.0409	0.0452	0.0460	0.0412	0.0446
	Adj. chi-square	0.0552	0.0482	0.0555	0.0533	0.0484	0.0549
	Proposed F I	0.0362	0.0346	0.0375	0.0336	0.0351	0.0342
	Proposed F II	0.0475	0.0407	0.0451	0.0453	0.0406	0.0453

Table 3: Actual sizes of the tests for $k = 4$ with generated covariates from $N(5, 1)$.

$n_1, \ldots, n_4$		$\sigma_1, \sigma_2, \sigma_3, \sigma_4$					
		1, 1, 1, 1	1,1.5,2,2.5	1,2,3,4	1,1,3,3	1,1,1,10	1,1,10,10
5,5,5,5	PB	0.0411	0.0463	0.0488	0.0519	0.0531	0.0545
	GPV	0.0164	0.0218	0.0228	0.0225	0.0230	0.0335
	Adj. chi-square	0.0238	0.0305	0.0316	0.0282	0.0300	0.0414
	Proposed F I	0.0163	0.0139	0.0236	0.0200	0.0202	0.0223
5,6,7,8	PB	0.0477	0.0454	0.0486	0.0473	0.0470	0.0497
	GPV	0.0249	0.0194	0.0250	0.0230	0.0244	0.0240
	Adj. chi-square	0.0352	0.0275	0.0358	0.0326	0.0330	0.0328
	Proposed F I	0.0172	0.0136	0.0224	0.0169	0.0178	0.0208
8,7,6,5	PB	0.0482	**0.0543**	**0.0548**	**0.0549**	0.0496	0.0502
	GPV	0.0204	0.0291	0.0378	0.0344	0.0348	0.0392
	Adj. chi-square	0.0300	0.0387	0.0476	0.0439	0.0423	0.0496
	Proposed F I	0.0166	0.0243	0.0306	0.0276	0.0277	0.0335
5,7,10,15	PB	0.0536	0.0491	0.0486	0.0486	0.0493	0.0502
	GPV	0.0360	0.0289	0.0290	0.0267	0.0253	0.0271
	Adj. chi-square	0.0419	0.0361	0.0364	0.0331	0.0303	0.0343
	Proposed F I	0.0215	0.0136	0.0168	0.0179	0.0156	0.0187
15,10,7,5	PB	0.0530	**0.0582**	**0.0579**	**0.0566**	0.0533	0.0551
	GPV	0.0282	0.0371	0.0448	0.0463	0.0420	0.0444
	Adj. chi-square	0.0359	0.0442	0.0526	0.0538	0.0473	0.0515
	Proposed F I	0.0199	0.0295	0.0355	0.0290	0.0283	0.0332
5,5,5,30	PB	**0.0574**	**0.0555**	**0.0550**	**0.0560**	0.0472	0.0532
	GPV	0.0264	0.0201	0.0211	0.0183	0.0157	0.0239
	Adj. chi-square	0.0337	0.0251	0.0263	0.0227	0.0191	0.0268
	Proposed F I	0.0181	0.0171	0.0153	0.0130	0.0121	0.0209
30,5 5,5	PB	**0.0567**	**0.0640**	**0.0609**	**0.0589**	**0.0539**	0.0547
	GPV	0.0251	0.0398	0.0363	0.0379	0.0264	0.0374
	Adj. chi-square	0.0326	0.0491	0.0464	0.0493	0.0339	0.0478
	Proposed F I	0.0179	0.0315	0.0345	0.0339	0.0221	0.0362
7,7,7,7	PB	0.0473	0.0495	0.0512	0.0498	0.0473	0.0493
	GPV	0.0267	0.0329	0.0312	0.0309	0.0366	0.0393
	Adj. chi-square	0.0397	0.0491	0.0458	0.0440	0.0515	0.0543
	Proposed F I	0.0201	0.0217	0.0294	0.0273	0.0277	0.0283
	Proposed F II	0.0228	0.0285	0.0275	0.0277	0.0320	0.0354
7,7,20,20	PB	0.0531	0.0511	0.0523	0.0503	0.0509	0.0498
	GPV	0.0382	0.0301	0.0363	0.0303	0.0326	0.0442
	Adj. chi-square	0.0437	0.0417	0.0400	0.0402	0.0350	0.0379
	Proposed F I	0.0266	0.0245	0.0234	0.0247	0.0206	0.0246
	Proposed F II	0.0265	0.0246	0.0236	0.0252	0.0214	0.0250
20,20,7,7	PB	0.0533	0.0538	0.0537	0.0533	0.0492	0.0509
	GPV	0.0371	0.0407	0.0459	0.0477	0.0393	0.0434
	Adj. chi-square	0.0499	0.0528	0.0590	0.0599	0.0506	0.0591
	Proposed F I	0.0312	0.0356	0.0354	0.0386	0.0316	0.0392
	Proposed F II	0.0281	0.0331	0.0377	0.0395	0.0315	0.0368
20,20 20,20	PB	0.0473	0.0494	0.0501	0.0501	0.0480	0.0504
	GPV	0.0433	0.0480	0.0455	0.0459	0.0403	0.0437
	Adj. chi-square	0.0498	0.0536	0.0513	0.0534	0.0462	0.0504
	Proposed F I	0.0369	0.0406	0.0436	0.0443	0.0352	0.0386
	Proposed F II	0.0427	0.0481	0.0453	0.0460	0.0398	0.0436

Table 4: Actual sizes of the tests for $k = 6$ with generated covariates from $N(5, 1)$.

		$\sigma_1, \sigma_2, \ldots, \sigma_6$					
$n_1, \ldots, n_6$		1,1,1 1,1,1	1,1,1.5 2,2,2.5	1,2,3 1,2,3	1,1,1 3,3,3	1,1,1 1,1,10	1,1,1 10,10,10
5,5,5,5,5,5	PB	0.0397	0.0442	0.0464	0.0488	0.0433	0.0509
	GPV	0.0151	0.0177	0.0199	0.0238	0.0155	0.0298
	Adj. chi-square	0.0300	0.0307	0.0328	0.0371	0.0267	0.0428
	Proposed F I	0.0102	0.0135	0.0154	0.0167	0.0130	0.0270
5,6,7,8,9,10	PB	0.0498	0.0470	0.0483	0.0472	0.0492	0.0487
	GPV	0.0275	0.0279	0.0273	0.0224	0.0213	0.0250
	Adj. chi-square	0.0436	0.0446	0.0403	0.0354	0.0354	0.0380
	Proposed F I	0.0168	0.0156	0.0150	0.0109	0.0133	0.0142
10,9,8,7,6,5	PB	0.0513	**0.0549**	**0.0554**	0.0539	0.0483	0.0526
	GPV	0.0256	0.0332	0.0301	0.0372	0.0345	0.0405
	Adj. chi-square	0.0421	0.0498	0.0456	0.0533	0.0461	0.0566
	Proposed F I	0.0138	0.0234	0.0217	0.0260	0.0211	0.0296
5,5,7,7,10,15	PB	0.0529	0.0460	0.0522	0.0470	0.0492	0.0491
	GPV	0.0256	0.0187	0.0225	0.0248	0.0191	0.0249
	Adj. chi-square	0.0403	0.0303	0.0337	0.0350	0.0315	0.0379
	Proposed F I	0.0157	0.0119	0.0155	0.0119	0.0116	0.0152
10,10,10,5,5,5	PB	0.0532	**0.0589**	**0.0577**	**0.0595**	0.0506	0.0527
	GPV	0.0229	0.0318	0.0252	0.0362	0.0281	0.0407
	Adj. chi-square	0.0321	0.0446	0.0368	0.0520	0.0404	0.0593
	Proposed F I	0.0124	0.0231	0.0185	0.0279	0.0167	0.0293
5,5,5,5,5,30	PB	0.0496	0.0461	0.0502	0.0540	0.0441	0.0502
	GPV	0.0216	0.0216	0.0173	0.0230	0.0116	0.0310
	Adj. chi-square	0.0365	0.0353	0.0307	0.0348	0.0195	0.0450
	Proposed F I	0.0151	0.0119	0.0117	0.0139	0.0111	0.0204
30,5,5,5,5,5	PB	0.0510	**0.0574**	**0.0575**	**0.0561**	0.0507	0.0535
	GPV	0.0197	0.0306	0.0319	0.0279	0.0215	0.0341
	Adj. chi-square	0.0341	0.0475	0.0505	0.0444	0.0339	0.0517
	Proposed F I	0.0131	0.0233	0.0226	0.0242	0.0154	0.0264
7,7,7,7,7,7	PB	0.0493	0.0488	0.0489	0.0489	0.0478	0.0497
	GPV	0.0251	0.0307	0.0321	0.0293	0.0266	0.0359
	Adj. chi-square	0.0425	0.0499	0.0543	0.0510	0.0439	0.0571
	Proposed F I	0.0155	0.0221	0.0188	0.0200	0.0189	0.0235
	Proposed F II	0.0193	0.0233	0.0250	0.0223	0.0212	0.0291
7,7,10,12,15,15	PB	0.0515	0.0493	0.0505	0.0498	0.0511	0.0482
	GPV	0.0344	0.0351	0.0372	0.0335	0.0288	0.0350
	Adj. chi-square	0.0510	0.0487	0.0527	0.0475	0.0436	0.0502
	Proposed F I	0.0239	0.0198	0.0217	0.0189	0.0207	0.0194
	Proposed F II	0.0244	0.0263	0.0274	0.0255	0.0224	0.0273
20,20,20,7,7,7	PB	0.0371	0.0407	0.0459	0.0477	0.0393	0.0434
	GPV	0.0499	0.0528	0.0590	0.0599	0.0506	0.0591
	Adj. chi-square	0.0312	0.0356	0.0354	0.0386	0.0316	0.0392
	Proposed F I	0.0281	0.0331	0.0377	0.0395	0.0315	0.0368
	Proposed F II	0.0533	0.0538	0.0537	0.0533	0.0492	0.0509
20,20,20 20,20,20	PB	0.0503	0.0498	0.0501	0.0495	0.0500	0.0507
	GPV	0.0478	0.0436	0.0454	0.0455	0.0413	0.0490
	Adj. chi-square	0.0548	0.0505	0.0553	0.0541	0.0490	0.0567
	Proposed F I	0.0350	0.0334	0.0347	0.0350	0.0321	0.0377
	Proposed F II	0.0470	0.0436	0.0450	0.0452	0.0406	0.0481

Table 5: The Adjusted nominal sizes and powers of the tests.

$\sigma_1 = 1, \sigma_2 = 1.5, \sigma_3 = 2, \sigma_4 = 2.5$								
		Adjusted	$\mu_1, \mu_2, \mu_3, \mu_4$					
	Actual	nominal	10,11	10,10.5	10,10	10,10	10,10	13,10
Test	size	size	12,13	11,11.5	11,11	12,12	12,14	10,10
PB	0.0490	0.0517	0.708	0.203	0.161	0.549	0.890	0.973
GPV	0.0288	0.0772	0.707	0.204	0.160	0.545	0.890	0.975
Adj. chi-square	0.0440	0.0602	0.707	0.201	0.161	0.545	0.888	0.974
Proposed F I	0.0247	0.0811	0.706	0.201	0.161	0.544	0.888	0.975
Proposed F II	0.0260	0.0810	0.707	0.201	0.161	0.545	0.888	0.974
$\sigma_1 = 1, \sigma_2 = 100, \sigma_3 = 0.1, \sigma_4 = 0.01$								
			10.2,10	10.5,10	10,10.1	10,10	10,10	10,10.5
			10,10	10,10	10.2,10.3	11,11	12,12	11,11.5
PB	0.0517	0.0487	0.061	0.116	0.377	0.345	0.893	1.000
GPV t	0.0499	0.0502	0.061	0.116	0.377	0.345	0.894	1.000
Adj. chi-square	0.0704	0.0271	0.062	0.117	0.378	0.346	0.894	1.000
Proposed F I	0.0427	0.0557	0.061	0.117	0.378	0.345	0.894	1.000
Proposed F II	0.0442	0.0544	0.062	0.117	0.378	0.345	0.894	1.000

4.3 Illustrative examples

In this section, we first illustrate the above tests by applying them to two sets of simulated data given by Ananda (1998) with four treatments and seven observations in each treatment group. It is seen that our simple tests produce essentially the same results as the more complicated tests. Note that in both these examples, the classical F test leads to the "wrong" conclusion. In addition, we present one example with a real data set.

Example 4.1. *The data were generated from an ANCOVA model with parameters*

$$\mu_1 = 10, \quad \mu_2 = 11, \quad \mu_3 = 12, \quad \mu_4 = 13, \quad \beta = 1, \quad \sigma_1 = 1, \quad \sigma_2 = 1.5, \quad \sigma_3 = 2, \quad \sigma_4 = 2.5.$$

For testing the equality of treatment effects, the results are given in Table 6. For this data, $m_G = 5.0$, $v_G = 56.62$. It is seen that the classical F test accepts the hypothesis of equality of μ_i's, whereas the other tests reject it.

Note that the p-values for our proposed F tests are slightly larger than the p-values of the other tests because these tests are "conservative" and thus are "less inclined" to reject the null hypothesis. In this particular example, the variances are not very unequal, and in such cases, our F tests are more conservative; see the entries for $k = 4$, $n_i = 7, 7, 7, 7$ in Tables 1 and 3.

Table 6: The p-values for the examples obtained based on different methods.

Example	classical F test	generalized F test	PB test	Adj. chi-square test	proposed F test I	proposed F test II
4.1	0.1719	0.0475	0.0506	0.0520	0.0753	0.0747
4.2	0.0335	0.2144	0.2135	0.2987	0.2267	0.2418
4.3	0.2722	0.1723	0.1454	0.2101	0.1992	0.1761

Example 4.2. *The data were generated from an ANCOVA model with parameters*

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 = 10, \quad \beta = 1, \quad \sigma_1 = 1, \quad \sigma_2 = 1.5, \quad \sigma_3 = 2, \quad \sigma_4 = 2.5.$$

For testing the equality of treatment effects, the results are given in Table 6. For this data, $m_G = 5.0$, $v_G = 56.6663$. *It is seen that the classical F test rejects the hypothesis of equality of* μ_i*'s, whereas the other tests accept it.*

Example 4.3. *To thoroughly examine the impact of diet on the maturation weight of guppy fish (Poecilia reticulata), a carefully controlled experiment was conducted involving three distinct groups of fish, each provided with a different diet designed to assess their growth outcomes. The resulting weights (y) of the guppies after the feeding period are documented in Table 16.2 from Morrison (1983), where they are presented alongside the corresponding initial weights (x) recorded before the dietary intervention. Each group consisted of seven observations, allowing for a comprehensive analysis of the effects of dietary variations on their growth patterns. For further details, readers are encouraged to refer to Rencher and Schaalje (2008), specifically to page 456, which discusses additional insights and statistical analyses. The results are presented in Table 6 for testing the equality of treatment effects among the three groups. It can be observed that all tests accept the hypothesis of equality of the* μ_i*'s.*

Acknowledgements

The authors are grateful to the anonymous referees for their helpful comments and suggestions to improve this manuscript.

References

Ananda, M.M. (1998). Bayesian and non-Bayesian solutions to analysis of covariance models under heteroscedasticity. *Journal of Econometrics*, **86**(1):177–192.

Arnold, S.F. (1981). *The Theory of Linear Models and Multivariate Analysis.* New York: Wiley.

Barrett, T.J. (2011). Computations using analysis of covariance. *Wiley Interdisciplinary Reviews: Computational Statistics*, **3**(3):260–268.

European Medicines Agency (2015). *Guideline on adjustment for baseline covariates in clinical trials.* Author London, England, EMA/CHMP/295050/2013.

Keselman, H.J., Huberty, C.J., Lix, L.M., Olejnik, S., Cribbie, R.A., Donahue, B., KowalchuK, R.K., Lowman, L.L., Petoskey, M.D., Keselman, J.C. and Levin, J.R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, **68**(3):350–386.

Khammar, A., Yarahmadi, M. and Madadizadeh, F. (2020). What is analysis of covariance (ANCOVA) and how to correctly report its results in medical research? *Iranian Journal of Public Health*, **49**(5):1016–1017.

Milliken, G.A. and Johnson, D.E. (2001). *Analysis of Messy Data, Volume III: Analysis of Covariance.* Volume 3, CRC Press.

Morrison, D.F. (1983). *Applied Linear Statistical Methods.* Prentice-Hall Englewood Cliffs, NJ.

R Core Team (2025). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rencher, A.C. and Schaalje, G.B. (2008). *Linear Models in Statistics.* New York: John Wiley & Sons.

Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM Approach.* John Wiley & Sons.

Sadooghi-Alvandi, S.M. and Jafari, A.A. (2013). A parametric bootstrap approach for one-way ANCOVA with unequal variances. *Communications in Statistics - Theory and Methods*, **42**(14):2473–2498.

Sadooghi-Alvandi, S.M., Jafari, A.A. and Mardani-Fard, H.A. (2012). One-way ANOVA with unequal variances. *Communications in Statistics-Theory and Methods*, **41**(22):4200–4221.

Tsui, K.W. and Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association*, **84**(406):602–607.

Welch, B.L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, **38**(3-4):330–336.

Appendix: Proofs

Let $\boldsymbol{q} = (q_1, \dots, q_k)'$, $\boldsymbol{w} = (w_1, \dots, w_k)'$, $\boldsymbol{a} = (a_1, \dots, a_k)'$, and $V = \left[\operatorname{diag}\left(\frac{s_i}{\sqrt{n_i}}\right)\right]$, where q_i, w_i, and a_i are as defined in (5). Note that $\boldsymbol{q}'\boldsymbol{q} = \sum_{i=1}^{k} w_i = 1$ and $A = \boldsymbol{a}'\boldsymbol{a}$. Then, it is easily verified that

$$\boldsymbol{a} = (I_k - \boldsymbol{q}q')\, V^{-1}\bar{\boldsymbol{x}}. \tag{7}$$

lemma 1. For any choice of a contrast matrix C, we have

$$VC'\big(CV^2C'\big)^{-1}CV = I_k - \boldsymbol{q}\boldsymbol{q}'.$$

lemma 2. For any choice of a contrast matrix C, we have

$$VC'\Big(C\hat{\Sigma}_{\text{obs}}C'\Big)^{-1}CV = I_k - \boldsymbol{q}\boldsymbol{q}' - K\boldsymbol{a}\boldsymbol{a}'. \tag{8}$$

Proof. We have

$$VC'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}CV = VC'\left(CV^2C' + \hat{\phi}_{\text{obs}}C\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'C'\right)^{-1}C = VC'\big(B + \boldsymbol{d}d'\big)^{-1}CV,$$

where $B = CV^2C'$ and $\boldsymbol{d} = \sqrt{\hat{\phi}_{\text{obs}}}C\bar{\boldsymbol{x}}$. Using the formula

$$\left(B + \boldsymbol{d}\boldsymbol{d}'\right)^{-1} = B^{-1} - \frac{1}{1 + \boldsymbol{d}'B^{-1}\boldsymbol{d}} B^{-1}\boldsymbol{d}\boldsymbol{d}'B^{-1},$$

(see Rencher and Schaalje, 2008), it follows that

$$VC'[B + \boldsymbol{d}\boldsymbol{d}']^{-1}CV = VC'B^{-1}CV - \frac{1}{1 + \boldsymbol{d}'B^{-1}\boldsymbol{d}} \left(VC'B^{-1}\boldsymbol{d}\right)\left(VC'B^{-1}\boldsymbol{d}\right)'. \tag{9}$$

But, using Lemma 1 and (7), it is easily verified that

$$\begin{aligned}
VC'B^{-1}CV &= VC'\left(CV^2C'\right)^{-1}CV = I_k - \boldsymbol{q}\boldsymbol{q}', \\
VC'B^{-1}\boldsymbol{d} &= \sqrt{\hat{\phi}_{\text{obs}}}VC'\left(CV^2C'\right)^{-1}C\bar{\boldsymbol{x}} = \sqrt{\hat{\phi}_{\text{obs}}}\left(I_k - \boldsymbol{q}\boldsymbol{q}'\right)\left(V^{-1}\bar{\boldsymbol{x}}\right) = \sqrt{\hat{\phi}_{\text{obs}}}\boldsymbol{a}, \\
\boldsymbol{d}'B^{-1}\boldsymbol{d} &= \hat{\phi}_{\text{obs}}(C\bar{\boldsymbol{x}})'\left(CV^2C'\right)^{-1}(C\bar{\boldsymbol{x}}) = \hat{\phi}_{\text{obs}}\left(V^{-1}\bar{\boldsymbol{x}}\right)'VC'\left(CV^2C'\right)^{-1}VC\left(V^{-1}\bar{\boldsymbol{x}}\right) \\
&= \hat{\phi}_{\text{obs}}\left(V^{-1}\bar{\boldsymbol{x}}\right)'\left(I_k - \boldsymbol{q}\boldsymbol{q}'\right)\left(V^{-1}\bar{\boldsymbol{x}}\right) = \hat{\phi}_{\text{obs}}\left(V^{-1}\bar{\boldsymbol{x}}\right)'\left(I_k - \boldsymbol{q}\boldsymbol{q}'\right)^2\left(V^{-1}\bar{\boldsymbol{x}}\right) \\
&= \hat{\phi}_{\text{obs}}\boldsymbol{a}'\boldsymbol{a} = \hat{\phi}_{\text{obs}}A.
\end{aligned}$$

Inserting these values in (9), (8) follows. □

Proof. **(Theorem 2.2):** It is sufficient to show that $T^* = T$. Let $\boldsymbol{D} = (D_1, \ldots, D_k)'$ and define $\boldsymbol{Z} = WC\boldsymbol{D}$. Then, it is easily verified that $\boldsymbol{D} \sim N_k(\boldsymbol{0}, \hat{\Sigma}_{\text{obs}})$ and, therefore $\boldsymbol{Z} \sim N_{k-1}(\boldsymbol{0}, I_{k-1})$, since $WC\hat{\Sigma}_{\text{obs}}C'W' = I$. Note that

$$\begin{aligned}
C'W'\boldsymbol{Z} &= C'W'WC\boldsymbol{D} = C'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}C\boldsymbol{D} = V^{-1}VC\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}VV^{-1}\boldsymbol{D} \\
&= V^{-1}\left(\boldsymbol{I} - \boldsymbol{q}\boldsymbol{q}' - K\boldsymbol{a}\boldsymbol{a}'\right)V^{-1}\boldsymbol{D}.
\end{aligned} \tag{10}$$

To simplify calculations, let $\boldsymbol{g} = (g_1, \ldots, g_k)'$, where $g_i = \frac{\sqrt{n_i}}{s_i}(D_i - \bar{D}_w)$. Then, it is easily verified that $\boldsymbol{g} = (I_k - \boldsymbol{q}\boldsymbol{q}')V^{-1}\boldsymbol{D}$, see (7). Then from (10), we have $C'W'\boldsymbol{Z} = V^{-1}[I - K\boldsymbol{a}\boldsymbol{a}']\boldsymbol{g}$, since

$$\boldsymbol{a}'V^{-1}\boldsymbol{D} = \bar{\boldsymbol{x}}'V^{-1}\left(I_k - \boldsymbol{q}\boldsymbol{q}'\right)V^{-1}\boldsymbol{D} = \bar{\boldsymbol{x}}'V^{-1}\left(I_k - \boldsymbol{q}\boldsymbol{q}'\right)^2V^{-1}\boldsymbol{D} = \boldsymbol{a}'\boldsymbol{g}.$$

Hence

$$\begin{aligned}
&\boldsymbol{Z}'WC\left[\operatorname{diag}\left(\frac{(n_i - 2)\,s_i^2}{n_iU_i}\right)\right]C'W'\boldsymbol{Z} \\
&= \boldsymbol{g}'\left[I - K\boldsymbol{a}\boldsymbol{a}'\right]V^{-1}\left[\operatorname{diag}\left(\frac{(n_i - 2)\,s_i^2}{n_iU_i}\right)\right]V^{-1}\left[I - K\boldsymbol{a}\boldsymbol{a}'\right]\boldsymbol{g} \\
&= \boldsymbol{g}'\left[I - K\boldsymbol{a}\boldsymbol{a}'\right]\left[\operatorname{diag}\left(\frac{n_i - 2}{U_i}\right)\right]\left[I - K\boldsymbol{a}\boldsymbol{a}'\right]\boldsymbol{g} \\
&= \boldsymbol{g}'\left[\operatorname{diag}\left(\frac{n_i - 2}{U_i}\right)\right]\boldsymbol{g} - 2K(\boldsymbol{a}'\boldsymbol{g})\boldsymbol{g}'\left[\operatorname{diag}\left(\frac{n_i - 2}{U_i}\right)\right]\boldsymbol{a}
\end{aligned}$$

$$
\begin{aligned}
&+K^2(\boldsymbol{a}'\boldsymbol{g})^2\boldsymbol{a}'\left[\operatorname{diag}\left(\frac{n_i-2}{U_i}\right)\right]\boldsymbol{a}\\
&=\sum_{i=1}^{k}\frac{n_i-2}{U_i}g_i^2-2K\left(\sum_{i=1}^{k}a_ig_i\right)\left(\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_ig_i\right)\\
&+K^2\left(\sum_{i=1}^{k}a_ig_i\right)^2\left(\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_i^2\right)\\
&=\sum_{i=1}^{k}\frac{n_i(n_i-2)}{s_i^2U_i}\left(D_i-\bar{D}_w\right)^2-2K\left(\sum_{i=1}^{k}\frac{\sqrt{n_i}(n_i-2)}{s_iU_i}a_i\left(D_i-\bar{D}_w\right)\right)\\
&\times\left(\sum_{i=1}^{k}\frac{\sqrt{n_i}}{s_i}a_i\left(D_i-\bar{D}_w\right)\right)+K^2\left\{\sum_{i=1}^{k}\frac{\sqrt{n_i}}{s_i}a_i\left(D_i-\bar{D}_w\right)\right\}^2\left[\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_i^2\right].
\end{aligned}\tag{11}
$$

Also, from (10) we have

$$
\begin{aligned}
\bar{\boldsymbol{x}}'C'W'\boldsymbol{Z}&=\bar{\boldsymbol{x}}'V^{-1}\left[(I_k-\boldsymbol{q}\boldsymbol{q}')-K\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)V^{-1}\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'V^{-1}\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)\right]V^{-1}\boldsymbol{D}\\
&=\boldsymbol{a}'\boldsymbol{g}-K\boldsymbol{a}'\boldsymbol{a}\boldsymbol{a}'\boldsymbol{g}=(\boldsymbol{a}'\boldsymbol{g})\left[1-K\left(\boldsymbol{a}'\boldsymbol{a}\right)\right]=H\left[1-KA\right],
\end{aligned}
$$

since $(I_k-\boldsymbol{q}\boldsymbol{q}')^2=(I_k-\boldsymbol{q}\boldsymbol{q}')$. Hence

$$
\begin{aligned}
\boldsymbol{Z}'WC\left[\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_is_i^2}{U_i}\right)\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'\right]C'W'\boldsymbol{Z}&=\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_is_i^2}{U_i}\right)\left(\bar{\boldsymbol{x}}'C'W'\boldsymbol{Z}\right)^2\\
&=\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_is_i^2}{U_i}\right)H^2\left(1-KA\right)^2.
\end{aligned}\tag{12}
$$

Combining (11) and (12), it follows that $T^*=T$, and the proof is completed. □

Proof. **(Theorem 2.4):** We use the well-known formula $E(T_G)=E[E(T_G|U_1,\ldots,U_k)]$. To find the conditional mean, T_G may be written in the form

$$T_G=\boldsymbol{Z}'WCMC'W'\boldsymbol{Z},$$

where $M=\left[\operatorname{diag}\left(\frac{(n_i-2)s_i^2}{n_iU_i}\right)\right]+\left(\sum_{i=1}^{k}\frac{(n_i-2)l_is_i^2}{U_i}\right)\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'$. Now recall that for a quadratic form $\boldsymbol{Z}'B\boldsymbol{Z}$, with $\boldsymbol{Z}\sim N(\boldsymbol{0},I)$, we have $E\left(\boldsymbol{Z}'B\boldsymbol{Z}\right)=\operatorname{trace}(B)$. Hence,

$$
\begin{aligned}
E\left(T_G|U_1,\ldots,U_k\right)&=\operatorname{trace}\left(WCMC'W'\right)=\operatorname{trace}\left(MC'W'WC\right)\\
&=\operatorname{trace}\left[MC'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}C\right]\\
&=\operatorname{trace}\left[MV^{-1}VC'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}CVV^{-1}\right]\\
&=\operatorname{trace}\left[MV^{-1}\left(I-\boldsymbol{q}\boldsymbol{q}'-K\boldsymbol{a}\boldsymbol{a}'\right)V^{-1}\right]\\
&=\operatorname{trace}\left(B_1\right)-K\operatorname{trace}\left(B_2\right)
\end{aligned}
$$

$$+\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_i s_i^2}{U_i}\right)\left[\operatorname{trace}(B_3)-K\operatorname{trace}(B_4)\right],$$

where

$$B_1=\left[\operatorname{diag}\left(\frac{(n_i-2)\,s_i^2}{n_iU_i}\right)\right]V^{-1}\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)V^{-1},\qquad B_3=\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'V^{-1}\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)V^{-1},$$

$$B_2=\left[\operatorname{diag}\left(\frac{(n_i-2)\,s_i^2}{n_iU_i}\right)\right]V^{-1}\boldsymbol{a}\boldsymbol{a}'V^{-1},\qquad B_4=\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'V^{-1}\boldsymbol{a}\boldsymbol{a}'V^{-1}.$$

But

$$\begin{aligned}
\operatorname{trace}(B_1) &= \operatorname{trace}\left(V^{-1}\left[\operatorname{diag}\left(\frac{(n_i-2)\,s_i^2}{n_iU_i}\right)\right]V^{-1}\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)\right)\\
&= \operatorname{trace}\left(\left[\operatorname{diag}\left(\frac{n_i-2}{U_i}\right)\right]\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)\right)=\sum_{i=1}^{k}\frac{n_i-2}{U_i}(1-w_i),\\
\operatorname{trace}(B_2) &= \operatorname{trace}\left\{\boldsymbol{a}'V^{-1}\left[\operatorname{diag}\left(\frac{(n_i-2)\,s_i^2}{n_iU_i}\right)\right]V^{-1}\boldsymbol{a}\right\}=\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_i^2,\\
\operatorname{trace}(B_3) &= \operatorname{trace}\left(\bar{\boldsymbol{x}}'V^{-1}\left(I_k-\boldsymbol{q}\boldsymbol{q}'\right)V^{-1}\bar{\boldsymbol{x}}\right)=\boldsymbol{a}'\boldsymbol{a}=A,\\
\operatorname{trace}(B_4) &= \bar{\boldsymbol{x}}'V^{-1}\boldsymbol{a}\boldsymbol{a}'V^{-1}\bar{\boldsymbol{x}}=\left(\bar{\boldsymbol{x}}'V^{-1}\boldsymbol{a}\right)^{\mathbf{2}}=\left(\boldsymbol{a}'\boldsymbol{a}\right)^{\mathbf{2}}=A^2,
\end{aligned}$$

since $\bar{\boldsymbol{x}}'V^{-1}\boldsymbol{a}=\bar{\boldsymbol{x}}'V^{-1}(\boldsymbol{I}-\boldsymbol{q}\boldsymbol{q}')V^{-1}\bar{\boldsymbol{x}}=\bar{\boldsymbol{x}}'V^{-1}\left(\boldsymbol{I}-\boldsymbol{q}\boldsymbol{q}'\right)^{\mathbf{2}}V^{-1}\bar{\boldsymbol{x}}=\boldsymbol{a}'\boldsymbol{a}$ Therefore,

$$\begin{aligned}
E\left(T_G|U_1,\ldots,U_k\right) &= \sum_{i=1}^{k}\frac{n_i-2}{U_i}(1-w_i)-K\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_i^2\\
&\quad+\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_i s_i^2}{U_i}\right)\left(A-KA^2\right).
\end{aligned}$$

Hence $E(T_G)=\sum_{i=1}^{k}E\left(\frac{1}{U_i}\right)(n_i-2)\left[1-w_i-Ka_i^2+A(1-AK)\,l_i s_i^2\right]$. Since $E\left(\frac{1}{U_i}\right)=\frac{1}{n_i-4}$. □

Proof. **(Theorem 2.6):** Using the same notation as in the proof of Theorem 2.4, we find the variance using the well-known formula $\operatorname{Var}(T_G)=E\left(\operatorname{Var}\left(T_G|U_1,\ldots,U_k\right)\right)+\operatorname{Var}\left(E\left(T_G|U_1,\ldots,U_k\right)\right)$. Recall that for a quadratic form $\boldsymbol{Z}'B\boldsymbol{Z}$ with $\boldsymbol{Z}\sim N(\mathbf{0},I)$, we have $\operatorname{Var}(\boldsymbol{Z}'B\boldsymbol{Z})=2\operatorname{trace}\left(B^2\right)$. Therefore,

$$\begin{aligned}
\operatorname{Var}\left(T_G|U_1,\ldots,U_k\right) &= 2\operatorname{trace}\left(WCMC'W'WCMC'W'\right)\\
&= 2\operatorname{trace}\left(MC'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}CMC'\left(C\hat{\Sigma}_{\text{obs}}C'\right)^{-1}C\right)\\
&= 2\operatorname{trace}\left[MV^{-1}\left(I-\boldsymbol{q}\boldsymbol{q}'-K\boldsymbol{a}\boldsymbol{a}'\right)V^{-1}MV^{-1}\right.\\
&\quad\left.\times\left(I-\boldsymbol{q}\boldsymbol{q}'-K\boldsymbol{a}\boldsymbol{a}'\right)V^{-1}\right],
\end{aligned}$$

After tedious calculations, we obtain

$$\begin{aligned}
\frac{1}{2}\mathrm{Var}\left(T_G|U_1,\ldots,U_k\right) &= \sum_{i=1}^{k}\frac{(n_i-2)^2}{U_i^2}\left(1-2w_i-2Ka_i^2\right)+\left(\sum_{i=1}^{k}\frac{n_i-2}{U_i}w_i\right)^2 \\
&+2K\left(\sum_{i=1}^{k}\frac{n_i-2}{U_i}q_ia_i\right)^2+K^2\left(\sum_{i=1}^{k}\frac{(n_i-2)}{U_i}a_i^2\right)^2 \\
&+2\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_is_i^2}{U_i}\right)\left(\sum_{i=1}^{k}\frac{n_i-2}{U_i}a_i^2\right)(1-KA)^2 \\
&+\left(\sum_{i=1}^{k}\frac{(n_i-2)\,l_is_i^2}{U_i}\right)^2A^2(1-KA)^2.
\end{aligned}$$

Note that $E(\frac{1}{U_i})=\frac{1}{n_i-4}$, $\mathrm{Var}(\frac{1}{U_i})=\frac{2}{(n_i-4)(n_i-6)}$, $E(\frac{1}{U_i^2})=\frac{1}{(n_i-4)(n_i-6)}$. Therefore,

$$\begin{aligned}
\frac{1}{2}E\left(\mathrm{Var}\left(T_G|U_1,\ldots,U_k\right)\right) &= \sum_{i=1}^{k}\frac{(n_i-2)^2}{(n_i-4)(n_i-6)}\left(1-2w_i-2Ka_i^2\right) \\
&+2\sum_{i=1}^{k}\frac{(n_i-2)^2}{(n_i-4)^2(n_i-6)}\left[w_i^2+2Kw_ia_i^2+K^2a_i^4\right. \\
&\left.+\,2(1-KA)^2l_is_i^2a_i^2+A^2(1-KA)^2l_i^2s_i^4\right] \\
&+\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}w_i\right)^2+2K\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}q_ia_i\right)^2 \\
&+K^2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}a_i^2\right)^2 \\
&+2(1-KA)^2\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}l_is_i^2\right)\left(\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}a_i^2\right) \\
&+A^2(1-KA)^2\sum_{i=1}^{k}\frac{n_i-2}{n_i-4}l_is_i^2. \qquad (13)
\end{aligned}$$

Also, from the proof of Theorem 2.4, we have

$$E\left(T_G|U_1,\ldots,U_k\right)=\sum_{i=1}^{k}\frac{n_i-2}{U_i}\left[1-w_i-Ka_i^2+A\left(1-AK\right)l_is_i^2\right].$$

Therefore,

$$\mathrm{Var}\left(E\left(T_G|U_1,\ldots,U_k\right)\right)=\sum_{i=1}^{k}\frac{2(n_i-2)^2}{(n_i-4)^2(n_i-6)}\left[1-w_i-Ka_i^2+A\left(1-AK\right)l_is_i^2\right]^2. \qquad (14)$$

Combining (13) and (14), the result follows. □