One-sample test for sub-mean vector in high-dimensional data

Riku Hosonuma[1*], Tamae Kawasaki[2] and Takashi Seo[3]

[1*]Department of Applied Mathematics, Graduate School of Science, Tokyo University of Science, Tokyo, Japan.
[2]Department of Economics, College of Economics, Aoyama Gakuin University, Tokyo, Japan.
[3]Department of Applied Mathematics, Faculty of Science, Tokyo University of Science, Tokyo, Japan.

Abstract

A Rao's $\boldsymbol{U}$-statistic is a classical hypotheses testing framework for sub-mean vectors based on Hotelling's $\boldsymbol{T^2}$ statistic. Therefore, it requires inverting the sample covariance matrix, and Consequently, Rao's $\boldsymbol{U}$ statistic is not applicable in settings where the dimensionality exceeds the sample size. In this study, we propose a high-dimensional analogue of Rao's $\boldsymbol{U}$-statistic by substituting Hotelling's $\boldsymbol{T^2}$ with the $\boldsymbol{L^2}$-norm-based statistic introduced by Zhang and Xu (2009). An approximate null distribution was derived using a two-cumulant matching method. Furthermore, estimators of the unknown parameters involved in the approximation were constructed using the unbiased estimators of the trace functionals of the covariance matrix. Finally, Monte Carlo simulations were conducted to investigate the finite-sample performance of the proposed procedure.

1 Introduction

This study focuses on the one-sample problem of hypothesis testing for a sub-mean vector, first discussed in Rao (1949), with Rao's U-statistic and its null distribution subsequently introduced in Siotani, Hayakawa, and Fujikoshi (1985). Various extensions of such hypothesis testing have been developed under different settings since Rao (1949). For instance, Rencher (2012) has introduced Rao's U-statistic for the two-sample problem, and Gupta, Xu, and Fujikoshi (2006) have extended it to

the k-sample case. Kawasaki, Naito, and Seo (2011) have proposed a Hotelling's T^2-type statistic for testing a sub-mean vector in the one-sample problem. Extensions to incomplete-data settings have also been investigated. Kawasaki and Seo (2016) have considered hypothesis testing for a sub-mean vector under two-step monotone missing data and developed a likelihood ratio test. Hosonuma, Kawasaki, and Seo (2025) have studied the one-sample sub-mean vector testing problem under two-step monotone missing data, and proposed a Rao's U-type statistic. More recently, Hosonuma, Kawasaki, and Seo (2026) have extended this framework to the two-sample problem.

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be distributed as the multivariate normal $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

We partitioned $\boldsymbol{x}_j$ into a $p_1 \times 1$ and $p_2 \times 1$ random vectors with $\boldsymbol{x}_j = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j})'$, where $\boldsymbol{x}_{ij} : p_i \times 1, i = 1, 2, j = 1, 2, \ldots, N$.

The sample mean vector and unbiased covariance matrix were defined as:

$$\begin{aligned} \overline{\boldsymbol{x}} =& \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{x}_j, \quad \overline{\boldsymbol{x}}_1 = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{x}_{1j}, \\ \boldsymbol{S} =& \frac{1}{n} \sum_{j=1}^{N} (\boldsymbol{x}_j - \overline{\boldsymbol{x}})(\boldsymbol{x}_j - \overline{\boldsymbol{x}})' = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix}, \end{aligned}$$

where $n = N - 1$ and $\boldsymbol{S}_{ij}$ is a $p_i \times p_j$matrix.

We considered the following sub-mean vector hypothesis:

$$H_0 : \boldsymbol{\mu}_2 = \boldsymbol{\mu}_{20} \text{ given } \boldsymbol{\mu}_1 = \boldsymbol{\mu}_{10} \text{ vs. } H_1 : \boldsymbol{\mu}_2 \neq \boldsymbol{\mu}_{20} \text{ given } \boldsymbol{\mu}_1 = \boldsymbol{\mu}_{10} \tag{1}$$

where $\boldsymbol{\mu}_{20}$ and $\boldsymbol{\mu}_{10}$ are known. Rao's U-statistic is defined as:

$$U = \frac{T_p^2 - T_{p_1}^2}{N - 1 + T_{p_1}^2},$$

where

$$\begin{aligned} T_p^2 =& N(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)' \boldsymbol{S}^{-1} (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0), \\ T_{p_1}^2 =& N(\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_{10})' \boldsymbol{S}_{11}^{-1} (\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_{10}), \end{aligned}$$

where T_p^2 and $T_{p_1}^2$ are the Hotelling's T^2-statistics. Under H_0, $(N - p)U/p_2$ is distributed as an F-distribution with p_2 and $N - p$ degrees of freedom. These results have been introduced by Siotani, Hayakawa, and Fujikoshi (1985).

However, when $p > N$ or $\boldsymbol{\Sigma}$ is singular, Hotelling's T^2 statistic is not well defined, and Rao's U-statistic cannot be applied. Dempster (1958, 1960) has proposed a non-exact test for high-dimensional settings to address this problem. In recent years, high-dimensional data are increasingly used in many scientific fields, and their dimensionity is comparable to or even higher than the sample size. Accordingly, various testing procedures have been developed for high-dimensional mean vectors, including the scale-invariant test of Srivastava and Du (2008) and the trace-based test of Schott (2007). Importantly research has also focused on evolving alternatives to Hotelling's T^2 test based on the L^2 norm. Bai and Saranadasa (1996) have proposed an L^2-norm-based test for the two-sample mean vector problem. Subsequently, Chen and Qin (2010) have developed an improved L^2-norm-based test, and Zhang and Xu (2009) have extended the L^2 norm approach to the one-sample mean vector problem. More recently, Zhang, Guo, Zhou, and Cheng (2020) and Zhang, Zhou, and Guo (2022) have advanced L^2-norm-based procedures for high-dimensional mean vector testing.

However, these studies have focused on testing the mean vectors, whereas hypothesis testing for a sub-mean vector in high-dimensional settings has not been investigated thus far. Accordingly, we considered the sub-mean vector hypothesis testing problem for high-dimensional data, and The propose a test statistic that preserves the structure of Rao's U-statistic while replacing Hotelling's T^2 statistic with the L^2-norm-based statistic of Zhang and Xu (2009). Specifically, we tested the following hypothesis:

$$H_0' : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \text{ vs. } H_1' : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \tag{2}$$

where $\boldsymbol{\mu}_0$ is known. For the one-sample problem (2), the L^2-norm-based test statistic can be specified as:

$$R_p = N||\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0||^2$$

Zhang and Xu (2009) imposed the following regular assumptions:

Assumption A.

1. $\dfrac{p}{N} \to \gamma \in (0, 1)$ as $N, p \to \infty$.
2. $\dfrac{\text{tr}(\boldsymbol{\Sigma})}{p} \to c_1 \in (0, \infty)$ as $p \to \infty$.
3. $\dfrac{\text{tr}(\boldsymbol{\Sigma}^2)}{p} \to c_2 \in (0, \infty)$ as $p \to \infty$.
4. $\dfrac{\lambda_{\max}}{\sqrt{p}} \to 0$ as $p \to \infty$,

where $\lambda_{\max} = \max_{1 \leq i \leq p} \lambda_i$, and $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of $\boldsymbol{\Sigma}$.

These assumptions are also similar to those imposed by Bai and Saranadasa (1996).

In Section 2, we propose a high-dimensional extension of Rao's U-statistic for testing a sub-mean vector and derive its approximate F-distribution. The finite-sample performance of the proposed procedure was investigated using Monte Carlo simulations, as detailed in Section 3. A numerical example is presented in Section 4, and concluding remarks are provided in Section 5.

2 Rao's U-statistic for high-dimensional data

We propose a test statistic for the sub-mean vector with high-dimensional data. In hypothesis (1), the Hotelling's T^2 statistic used for Rao's U-statistic is replaced by a L^2-norm-based test statistic to yield a new test statistic, U_H, as defined below.

$$U_H = \frac{R_p - R_{p_1}}{N - 1 + R_{p_1}},$$

where

$$\begin{aligned} R_p =& N||\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0||^2, \\ R_{p_1} =& N||\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_{10}||^2. \end{aligned}$$

We impose the following assumptions to derive the asymptotic null distribution of the proposed statistic.

Assumption B.

1. $\frac{p_1}{N} \to \gamma_1 \in (0,1)$ and $\frac{p_2}{N} \to \gamma_2 \in (0,1)$, as $N, p_1, p_2 \to \infty$.
2. $\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{11})}{p_1} \to c_1^{(1)} \in (0,\infty)$ and $\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{22})}{p_2} \to c_1^{(2)} \in (0,\infty)$, as $p_1, p_2 \to \infty$.
3. $\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{11}^2)}{p_1} \to c_2^{(1)} \in (0,\infty)$ and $\frac{\mathrm{tr}(\boldsymbol{\Sigma}_{22}^2)}{p_2} \to c_2^{(2)} \in (0,\infty)$, as $p_1, p_2 \to \infty$.
4. $\frac{\lambda_{\max}^{(1)}}{\sqrt{p_1}} \to 0$ and $\frac{\lambda_{\max}^{(2)}}{\sqrt{p_2}} \to 0$, as $p_1, p_2 \to \infty$,

where $\lambda_1^{(1)}, \ldots, \lambda_{p_1}^{(1)}$ and $\lambda_1^{(2)}, \ldots, \lambda_{p_2}^{(2)}$ denote the eigenvalues of $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$, respectively, with $\lambda_{\max}^{(1)} = \max_{1 \le i \le p_1} \lambda_i^{(1)}$ and $\lambda_{\max}^{(2)} = \max_{1 \le i \le p_2} \lambda_i^{(2)}$.

We derived the asymptotic null distribution of U_H to test the hypothesis described in (1). Let $\overset{d}{=}$ and $\overset{d}{\approx}$ denote the equality and approximate equality in distribution, respectively, and let χ_d^2 denote a chi-squared distribution with d degrees of freedom.

Zhang and Xu (2009) have shown that

$$R_p \overset{d}{=} \sum_{r=1}^{p} \lambda_r A_r, \quad A_r \sim \chi_1^2.$$

Similarly,

$$R_{p_1} \overset{d}{=} \sum_{r=1}^{p_1} \lambda_r A_r, \quad A_r \sim \chi_1^2.$$

Hence,

$$U_H = \frac{R_p - R_{p_1}}{N - 1 + R_{p_1}}$$

$$\stackrel{d}{=} \frac{\sum_{r=1}^{p} \lambda_r A_r - \sum_{r=1}^{p_1} \lambda_r A_r}{N - 1 + \sum_{r=1}^{p_1} \lambda_r A_r}$$
$$= \frac{\sum_{r=p_1+1}^{p} \lambda_r A_r}{N - 1 + \sum_{r=1}^{p_1} \lambda_r A_r}.$$

We approximated the numerator and denominator of U_H separately using scaled chi-square distributions to obtain its approximate null distribution, via a two-cumulant matching method.

First, we considered the denominator. Matching the first two cumulants yields

$$N - 1 + \sum_{r=1}^{p_1} \lambda_r A_r \stackrel{d}{\approx} \beta_0 \chi^2_{d_0},$$

where

$$\beta_0 = \frac{\mathrm{tr}(\mathbf{\Sigma}_{11}^2)}{N - 1 + \mathrm{tr}(\mathbf{\Sigma}_{11})}, \quad d_0 = \frac{\{N - 1 + \mathrm{tr}(\mathbf{\Sigma}_{11})\}^2}{\mathrm{tr}(\mathbf{\Sigma}_{11}^2)}.$$

Similarly, for the numerator,

$$\sum_{r=p_1+1}^{p} \lambda_r A_r \stackrel{d}{\approx} \beta_1 \chi^2_{d_1},$$

where

$$\beta_1 = \frac{\mathrm{tr}(\mathbf{\Sigma}_{22}^2)}{\mathrm{tr}(\mathbf{\Sigma}_{22})}, \quad d_1 = \frac{\mathrm{tr}^2(\mathbf{\Sigma}_{22})}{\mathrm{tr}(\mathbf{\Sigma}_{22}^2)}.$$

Consequently,

$$U_H \stackrel{d}{\approx} \frac{\beta_1 \chi^2_{d_1}}{\beta_0 \chi^2_{d_0}}.$$

After a simple algebraic manipulation, we obtained

$$\frac{\beta_0 d_0}{\beta_1 d_1} U_H \stackrel{d}{\approx} \frac{\chi^2_{d_1}/d_1}{\chi^2_{d_0}/d_0} = F_{d_1, d_0}.$$

Therefore, under the null hypothesis, the proposed statistic was approximately distributed as F_{d_1,d_0}. Consequently, the proposed test could be implemented by employing this approximate F-distribution as the reference.

This approximation depends on the unknown parameters β_0, β_1, d_0, and d_1, which are functions of $\mathrm{tr}(\mathbf{\Sigma}_{11})$, $\mathrm{tr}(\mathbf{\Sigma}_{22})$, $\mathrm{tr}(\mathbf{\Sigma}_{11}^2)$, and $\mathrm{tr}(\mathbf{\Sigma}_{22}^2)$. These quantities were unknown in practice, and had to be estimated from the data. Estimating the trace functionals of the covariance matrices is central to high-dimensional statistics. Yamada and Himeno (2015, 2023) and Yamada, Himeno, Tillander, and Pavlenko (2023) have investigated estimating such quantities without bias.

Let

$$\boldsymbol{S} = \frac{1}{N-1} \sum_{i=1}^{N} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})' = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{pmatrix}$$

denote the sample covariance matrix. A straightforward approach involves estimating these trace functionals by replacing $\boldsymbol{\Sigma}$ with $\boldsymbol{S}$. However, such plug-in estimators are generally biased, particularly for $\mathrm{tr}^2(\boldsymbol{\Sigma})$ and $\mathrm{tr}(\boldsymbol{\Sigma}^2)$, in high-dimensional settings. Therefore, following Zhang and Xu (2009), we employed the unbiased estimators

$$\widehat{\mathrm{tr}^2(\boldsymbol{\Sigma})} = \frac{(N-1)N}{(N-2)(N+1)}\left[\mathrm{tr}^2(\boldsymbol{S}) - \frac{2}{N}\mathrm{tr}(\boldsymbol{S}^2)\right],$$
$$\widehat{\mathrm{tr}(\boldsymbol{\Sigma}^2)} = \frac{(N-1)^2}{(N-2)(N+1)}\left[\mathrm{tr}(\boldsymbol{S}^2) - \frac{1}{N-1}\mathrm{tr}^2(\boldsymbol{S})\right].$$

Applying these to $\boldsymbol{S}_{11}$ and $\boldsymbol{S}_{22}$ yields $\widehat{\mathrm{tr}^2(\boldsymbol{\Sigma}_{ii})}$ and $\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_{ii}^2)}$ $(i = 1, 2)$, respectively. Moreover, we estimated $\mathrm{tr}(\boldsymbol{\Sigma}_{11})$ and $\mathrm{tr}(\boldsymbol{\Sigma}_{22})$ by $\mathrm{tr}(\boldsymbol{S}_{11})$ and $\mathrm{tr}(\boldsymbol{S}_{22})$, respectively.

Substituting these estimators into the expressions for β_0, β_1, d_0, and d_1, we obtained

$$\hat{\beta}_0 = \frac{\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_{11}^2)}}{N - 1 + \mathrm{tr}(\boldsymbol{S}_{11})}, \quad \hat{d}_0 = \frac{(N-1)^2 + 2(N-1)\mathrm{tr}(\boldsymbol{S}_{11}) + \widehat{\mathrm{tr}^2(\boldsymbol{\Sigma}_{11})}}{\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_{11}^2)}},$$

and

$$\hat{\beta}_1 = \frac{\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_{22}^2)}}{\mathrm{tr}(\boldsymbol{S}_{22})}, \quad \hat{d}_1 = \frac{\widehat{\mathrm{tr}^2(\boldsymbol{\Sigma}_{22})}}{\widehat{\mathrm{tr}(\boldsymbol{\Sigma}_{22}^2)}}.$$

Hence, under the null hypothesis,

$$\frac{\hat{\beta}_0\hat{d}_0}{\hat{\beta}_1\hat{d}_1}U_H \overset{d}{\approx} F_{\hat{d}_1,\hat{d}_0}.$$

Accordingly, the proposed test rejected H_0 at significance level α whenever

$$\frac{\hat{\beta}_0\hat{d}_0}{\hat{\beta}_1\hat{d}_1}U_H > F_\alpha(\hat{d}_1, \hat{d}_0),$$

where $F_\alpha(\hat{d}_1, \hat{d}_0)$ denotes the upper 100α percentile of the F-distribution with degrees of freedom $\hat{d}_1$ and $\hat{d}_0$.

3 Monte Carlo Simulations

We performed Monte Carlo simulations with 10^6 replications to evaluate the finite-sample performance of Rao's U-statistic, the proposed high-dimensional statistic U_H, and its bias-corrected version.

The empirical Type I error rates are defined as

$$P_U = \Pr\left\{\frac{N-p}{p_2}U > F_{p_2,N-p}(\alpha)\right\},$$

$$P_H = \Pr\left\{ \frac{\tilde{\beta}_0 \tilde{d}_0}{\tilde{\beta}_1 \tilde{d}_1} U_H > F_{\tilde{d}_1, \tilde{d}_0}(\alpha) \right\},$$

$$P_{BH} = \Pr\left\{ \frac{\hat{\beta}_0 \hat{d}_0}{\hat{\beta}_1 \hat{d}_1} U_H > F_{\hat{d}_1, \hat{d}_0}(\alpha) \right\},$$

where $F_{k,l}(\alpha)$ denotes the upper 100α percentile of the F-distribution with degrees of freedom k and l. $\tilde{\beta}_0$, $\tilde{\beta}_1$, $\tilde{d}_0$, and $\tilde{d}_1$ were obtained using the plug-in estimators, with $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ replaced by $\boldsymbol{S}_{11}$ and $\boldsymbol{S}_{22}$, respectively, in the expressions for β_0, β_1, d_0, and d_1. By contrast, $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{d}_0$, and $\hat{d}_1$ were based on the bias-corrected estimators introduced in Section 2.

We conducted Monte Carlo simulations with 10^6 replications under several combinations of (p_1, p_2, N) to investigate the impact of the dimensionality and the sample size on the finite-sample performance of the proposed procedure. Specifically, the partition ratio $p_1 : p_2$ was assumed as $1 : 1$, $1 : 2$, $1 : 3$, and $2 : 1$. For each ratio, p_1 ranged from $4 - 60$ in increments of 2, with p_2 determined accordingly. The sample size was set to $N = p_1$, $2p_1$, $3p_1$, and $p_1/2$. Under the null hypothesis, random samples were generated independently from the p-variate normal distribution $N_p(\mathbf{0}, \boldsymbol{I}_p)$, where $p = p_1 + p_2$.

The empirical Type I error rate at the nominal significance levels $\alpha = 0.05$ is summarized in Tables $1 - 4$, which correspond to the partition ratios $p_1 : p_2 = 1 : 1$, $1 : 2$, $1 : 3$, and $2 : 1$, respectively.

Tables $1 - 4$ show that when $N > p$, the Type I error rate of Rao's U-statistic, P_U, is closest to the nominal significance level. This outcome is expected because Rao's U-statistic follows the F-distribution under the null hypothesis in this setting.

By contrast, the Type I error rate of the proposed statistic without the bias correction, P_H, does not satisfactorily control the Type I error rate. However, the Type I error rate P_{BH} was substantially improved by employing the bias-corrected estimators of the trace functionals, and remained close to the nominal significance level even when $p \geq N$, where Rao's U-statistic is no longer applicable. Although Tables $1 - 4$ only report the finite-dimensional settings, additional simulation results (not reported here) confirm that the empirical Type I error rate of the bias-corrected procedure, P_{BH}, converges to the nominal significance level as the sample size and dimensionality increase.

These results indicate that the proposed bias-corrected procedure enables hypothesis testing in high-dimensional settings with $p \geq N$ over a wide range of sample sizes and dimensional configurations. Although similar tendencies were observed for all partition ratios, the empirical Type I error rate of P_{BH} became slightly conservative when the difference between p_1 and p_2 was relatively large, e.g., $p_1 : p_2 = 1 : 3$.

4 Numerical Example

In this section, we illustrate the proposed test using the white wine data set described in Cortez (2009). The original dataset contained $4,898$ observations. All variables were transformed using the Box–Cox transformation. We randomly selected 20 and

10 observations from the transformed data set, respectively, to demonstrate both the scenarios wherein $N > p$ and $N = p$..

The variables were arranged in the following order: fixed acidity, citric acid, pH, volatile acidity, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, sulphates, and alcohol. As the fixed acidity, citric acid, and pH we determined by the characteristics of the grapes used in producing the wine, the test was performed under the assumption that their mean vector is given. Therefore, we have $N = 20, p = 10, p_1 = 3$, and $p_2 = 7$.

We considered the following hypothesis for the data after the Box–Cox transformation:

$$\begin{aligned} H_0 : \ \boldsymbol{\mu}_2 = (-1.646,\ 1.233,\ -7.312,\ 8.679,\ -0.007,\ -0.896,\ 0.543)' \\ \text{given } \boldsymbol{\mu}_1 = (0.701,\ 0.293,\ 0.553)' \\ \text{vs.} \quad H_1 : \ \boldsymbol{\mu}_2 = (-1.646,\ 1.233,\ -7.312,\ 8.679,\ -0.007,\ -0.896,\ 0.543)' \\ \text{given } \boldsymbol{\mu}_1 = (0.701,\ 0.293,\ 0.553)' \end{aligned}$$

Subsequently, we computed

$$\frac{N-p}{p_2} U = 4.828,$$

for Rao's U-statistic,

$$\frac{\tilde{\beta}_0 \tilde{d}_0}{\tilde{\beta}_1 \tilde{d}_1} U_H = 2.889,$$

for the proposed statistic, and

$$\frac{\hat{\beta}_0 \hat{d}_0}{\hat{\beta}_1 \hat{d}_1} U_H = 3.027,$$

for the bias-corrected version of the proposed statistic.

The corresponding critical values were $F_{p_2,N-p}(0.05) = 3.135$ and $F_{p_2,N-p}(0.01) = 5.200$, $F_{\tilde{d}_1,\tilde{d}_0}(0.05) = 3.014$ and $F_{\tilde{d}_1,\tilde{d}_0}(0.01) = 4.646$, and $F_{\hat{d}_1,\hat{d}_0}(0.05) = 2.901$ and $F_{\hat{d}_1,\hat{d}_0}(0.01) = 4.401$, respectively. For the bias-corrected version of the proposed statistic, the null hypothesis was rejected at the 0.05 significance level but not at the 0.01 level, which agrees with the result obtained using Rao's U-statistic. By contrast, the proposed statistic (without bias correction) did not lead to the same testing result, indicating that the bias correction is essential.

Next, we considered the case $N = p = 10$, while keeping the remaining settings unchanged. Since $N = p$, Rao's U-statistic cannot be applied because the denominator degrees of freedom of the corresponding F-distribution become zero.

For the proposed statistic, we obtained

$$\frac{\tilde{\beta}_0\tilde{d}_0}{\tilde{\beta}_1\tilde{d}_1}U_H = 2.886,$$

and, after bias correction,

$$\frac{\hat{\beta}_0\hat{d}_0}{\hat{\beta}_1\hat{d}_1}U_H = 3.241.$$

The corresponding critical values were $F_{\tilde{d}_1,\tilde{d}_0}(0.05) = 3.285$, $F_{\tilde{d}_1,\tilde{d}_0}(0.01) = 5.258$, $F_{\hat{d}_1,\hat{d}_0}(0.05) = 3.108$, and $F_{\hat{d}_1,\hat{d}_0}(0.01) = 4.853$. For the bias-corrected version of the proposed statistic, the null hypothesis was rejected at the 0.05 significance level but not at the 0.01 level, which is identical the testing conclusion obtained in the $N = 20$ case. Although Rao's U-statistic is not applicable when $N = p$, the proposed procedure could still be performed successfully and yielded the same testing conclusion as in the $N > p$ case.

These numerical examples demonstrate two important features of the proposed procedure. First, when $N > p$, the bias-corrected version of the proposed statistic yielded the same testing conclusion as Rao's U-statistic, whereas the uncorrected version could lead to a different conclusion, thus highlighting the necessity of bias correction. Second, when $N = p$, Rao's U-statistic was no longer applicable, whereas the proposed statistic remained valid, and provided the same testing conclusion as for the $N > p$ case.

5 Conclusions

In this study, we considered the one-sample problem of testing a sub-mean vector in high-dimensional settings. We proposed a high-dimensional extension of Rao's U-statistic and derived its approximate F-distribution using scaled chi-square approximations. Bias-corrected estimators were introduced for the trace functionals to improve the approximation accuracy.

Through Monte Carlo simulations, we demonstrated that the bias-corrected procedure U_H satisfactorily controlled the Type I error rate even when $p \geq N$, where in Rao's U-statistic is no longer applicable. Finally, a numerical example using a white wine dataset illustrated that the proposed procedure yields the same testing conclusion as Rao's U-statistic when $N > p$ and remains applicable when $N = p$. These results indicate that the proposed bias-corrected extension of Rao's U-statistic is a useful and robust tool for testing sub-mean vectors in high-dimensional settings.

References

[1] Bai, Z. D. and Saranadasa, H. (1996). Effect of high dimension: By an example of a two sample problem. *Statist. Sinica*, **6**, 311-329.

[2] Chen, S. X. and Qin, Y.-L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Ann. Statist.*, **38**, 808-835.

[3] Cortez, P., Cerdeira, A., Almeida, F., Matos, T., and Reis, J. (2009). Modeling wine preference by data mining from physicochemical properties. *Decis. Support Syst.*, **47**, 547-553.

[4] Dempster, A. P. (1958). A high dimensional two sample significance test. *Ann. Math. Statist.*, **29**, 995-1010.

[5] Dempster, A. P. (1960). A significance test for the separation of two highly multivariate small samples. *Biometrics*, **16**, 41-50.

[6] Gupta, A. K., Xu, J., and Fujikoshi, Y. (2006). An asymptotic expansion of the distribution of Rao's U-statistic under a general condition. *J. Multivariate Anal.*, **97**, 492-513.

[7] Hosonuma, R., Kawasaki, T., and Seo, T. (2025). On the extension of test statistics for the sub-mean vector under two-step monotone missing data. *Sankhyā B*, **87**, 434-467.

[8] Hosonuma, R., Kawasaki, T., and Seo, T. (2026). Two-sample tests of sub-mean vectors under two-step monotone missing data. *Int. J. Stat. Probab.*, **15**, 83-93.

[9] Kawasaki, T., Naito, K., and Seo, T. (2011). T^2 type test statistic and simultaneous confidence intervals for sub-mean vectors in k-sample problem. *Int. J. Stat. Probab.*, **9**, 1-8.

[10] Kawasaki, T. and Seo, T. (2016). A test for subvector of mean vector with two-step monotone missing data. *SUT J. Math.*, **52**, 21-39.

[11] Rao, C. R. (1949). On some problems arising out of discrimination with multiple characters. *Sankhyā*, **9**, 343–366.

[12] Rencher, A. C., and Christensen, W. F. (2012). *Methods of Multivariate Analysis*, 2nd ed. Hoboken, NJ: Wiley.

[13] Schott, J. R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Comput. Statist. Data Anal.*, **51**, 6535–6542.

[14] Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. Ohio: American Sciences Press.

[15] Srivastava, M. S. and Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *J. Multivariate Anal.*, **99**, 386-402.

[16] Yamada, T. and Himeno, T. (2015). Testing homogeneity of mean vectors under heteroscedasticity in high-dimension. *J. Multivariate Anal.*, **139**, 7-27.

[17] Yamada, T. and Himeno, T. (2023). High-dimensional asymptotic expansion of the null distribution for $L2$ norm based MANOVA testing statistic under general distribution. *J. Statist. Plann. Inference*, **224**, 9–26.

[18] Yamada, T., Himeno, T., Tillander, A., and Pavlenko, T. (2023). Test for mean matrix in GMANOVA model under heteroscedasticity and non-normality for high-dimensional data. *Theory Probab. Math. Statist.*, **109**, 129–158.

[19] Zhang, J. and Xu, J. (2009). On the k-sample Behrens-Fisher problem for high-dimensional data. *Sci. China Ser. A*, **52**, 1285–1304.

[20] Zhang, J., Guo, B., Zhou, B., and Cheng, C. (2020). A simple two-sample test in high dimensions based on L^2-norm. *J. Amer. Statist. Assoc.*, **115**, 1011–1027.

[21] Zhang, J., Zhou, B., and Guo, B. (2022). Testing high-dimensional mean vector with applications: a normal reference approach.. *Statist. Papers*, **63**, 1105–1137.

Table 1: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 1$, and $\alpha = 0.05$

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	4	12	6.40	0.67	0.050	0.042	0.073
6	6	18	4.30	0.57	0.050	0.036	0.062
8	8	24	3.45	0.52	0.050	0.034	0.058
10	10	30	2.98	0.48	0.050	0.033	0.056
12	12	36	2.69	0.46	0.050	0.032	0.055
14	14	42	2.48	0.44	0.050	0.032	0.053
16	16	48	2.33	0.43	0.050	0.031	0.053
18	18	54	2.22	0.42	0.050	0.031	0.052
20	20	60	2.13	0.41	0.050	0.031	0.052
22	22	66	2.05	0.40	0.050	0.031	0.052
24	24	72	1.98	0.39	0.050	0.031	0.052
26	26	78	1.93	0.38	0.050	0.031	0.052
28	28	84	1.88	0.38	0.050	0.030	0.051
30	30	90	1.84	0.37	0.050	0.030	0.051
40	40	120	1.69	0.36	0.050	0.030	0.051
50	50	150	1.60	0.34	0.050	0.030	0.051
60	60	180	1.53	0.33	0.050	0.029	0.050
4	4	8	nan	0.95	0.000	0.036	0.084
6	6	12	nan	0.80	0.000	0.030	0.067
8	8	16	nan	0.72	0.000	0.028	0.062
10	10	20	nan	0.67	0.000	0.026	0.058
12	12	24	nan	0.63	0.000	0.026	0.056
14	14	28	nan	0.60	0.000	0.025	0.055
16	16	32	nan	0.58	0.000	0.025	0.054
18	18	36	nan	0.57	0.000	0.025	0.054
20	20	40	nan	0.55	0.000	0.024	0.053
22	22	44	nan	0.54	0.000	0.024	0.052
24	24	48	nan	0.53	0.000	0.024	0.052
26	26	52	nan	0.52	0.000	0.024	0.052
28	28	56	nan	0.51	0.000	0.024	0.052
30	30	60	nan	0.51	0.000	0.024	0.052
40	40	80	nan	0.48	0.000	0.023	0.051
50	50	100	nan	0.46	0.000	0.023	0.051
60	60	120	nan	0.45	0.000	0.023	0.050

Table 1: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 1$, and $\alpha = 0.05$ (continued)

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	4	4	nan	1.72	0.000	0.020	0.117
6	6	6	nan	1.36	0.000	0.014	0.084
8	8	8	nan	1.19	0.000	0.013	0.071
10	10	10	nan	1.09	0.000	0.012	0.065
12	12	12	nan	1.02	0.000	0.012	0.061
14	14	14	nan	0.97	0.000	0.011	0.059
16	16	16	nan	0.93	0.000	0.011	0.057
18	18	18	nan	0.90	0.000	0.011	0.056
20	20	20	nan	0.87	0.000	0.011	0.055
22	22	22	nan	0.85	0.000	0.011	0.055
24	24	24	nan	0.83	0.000	0.011	0.054
26	26	26	nan	0.81	0.000	0.011	0.054
28	28	28	nan	0.80	0.000	0.011	0.053
30	30	30	nan	0.79	0.000	0.011	0.053
40	40	40	nan	0.74	0.000	0.011	0.052
50	50	50	nan	0.71	0.000	0.011	0.051
60	60	60	nan	0.69	0.000	0.010	0.051
4	4	2	nan	3.13	0.000	0.008	0.000
6	6	3	nan	2.21	0.000	0.003	0.136
8	8	4	nan	1.84	0.000	0.002	0.099
10	10	5	nan	1.64	0.000	0.002	0.083
12	12	6	nan	1.51	0.000	0.002	0.075
14	14	7	nan	1.42	0.000	0.002	0.070
16	16	8	nan	1.35	0.000	0.002	0.066
18	18	9	nan	1.29	0.000	0.002	0.064
20	20	10	nan	1.25	0.000	0.002	0.062
22	22	11	nan	1.21	0.000	0.002	0.060
24	24	12	nan	1.18	0.000	0.002	0.059
26	26	13	nan	1.15	0.000	0.002	0.058
28	28	14	nan	1.13	0.000	0.002	0.057
30	30	15	nan	1.11	0.000	0.002	0.057
40	40	20	nan	1.04	0.000	0.002	0.055
42	42	21	nan	1.02	0.000	0.002	0.054
44	44	22	nan	1.01	0.000	0.002	0.054
46	46	23	nan	1.00	0.000	0.002	0.054
48	48	24	nan	1.00	0.000	0.002	0.053
50	50	25	nan	0.99	0.000	0.002	0.054
52	52	26	nan	0.98	0.000	0.002	0.053
54	54	27	nan	0.97	0.000	0.002	0.053
56	56	28	nan	0.97	0.000	0.002	0.053
58	58	29	nan	0.96	0.000	0.002	0.053
60	60	30	nan	0.95	0.000	0.002	0.053

Table 2: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 2$, and $\alpha = 0.05$

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	8	12	nan	1.10	0.000	0.025	0.065
6	12	18	nan	0.96	0.000	0.023	0.058
8	16	24	nan	0.89	0.000	0.022	0.055
10	20	30	nan	0.84	0.000	0.021	0.054
12	24	36	nan	0.80	0.000	0.020	0.053
14	28	42	nan	0.78	0.000	0.020	0.052
16	32	48	nan	0.75	0.000	0.020	0.052
18	36	54	nan	0.74	0.000	0.020	0.051
20	40	60	nan	0.72	0.000	0.020	0.051
22	44	66	nan	0.71	0.000	0.020	0.051
24	48	72	nan	0.70	0.000	0.020	0.051
26	52	78	nan	0.69	0.000	0.019	0.051
28	56	84	nan	0.69	0.000	0.020	0.051
30	60	90	nan	0.68	0.000	0.019	0.051
40	80	120	nan	0.65	0.000	0.019	0.050
50	100	150	nan	0.63	0.000	0.019	0.051
60	120	180	nan	0.62	0.000	0.019	0.050
4	8	8	nan	1.59	0.000	0.017	0.069
6	12	12	nan	1.35	0.000	0.015	0.060
8	16	16	nan	1.23	0.000	0.014	0.056
10	20	20	nan	1.16	0.000	0.014	0.055
12	24	24	nan	1.11	0.000	0.013	0.053
14	28	28	nan	1.07	0.000	0.014	0.053
16	32	32	nan	1.03	0.000	0.013	0.052
18	36	36	nan	1.01	0.000	0.013	0.051
20	40	40	nan	0.99	0.000	0.013	0.051
22	44	44	nan	0.97	0.000	0.013	0.051
24	48	48	nan	0.96	0.000	0.013	0.051
26	52	52	nan	0.94	0.000	0.013	0.051
28	56	56	nan	0.93	0.000	0.013	0.051
30	60	60	nan	0.92	0.000	0.013	0.051
40	80	80	nan	0.88	0.000	0.013	0.050
50	100	100	nan	0.86	0.000	0.012	0.050
60	120	120	nan	0.84	0.000	0.012	0.050

Table 2: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 2$, and $\alpha = 0.05$ (continued)

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	8	4	nan	2.93	0.000	0.005	0.088
6	12	6	nan	2.35	0.000	0.004	0.067
8	16	8	nan	2.08	0.000	0.004	0.060
10	20	10	nan	1.92	0.000	0.004	0.057
12	24	12	nan	1.81	0.000	0.004	0.055
14	28	14	nan	1.73	0.000	0.004	0.054
16	32	16	nan	1.67	0.000	0.004	0.053
18	36	18	nan	1.62	0.000	0.004	0.053
20	40	20	nan	1.58	0.000	0.004	0.052
22	44	22	nan	1.54	0.000	0.004	0.052
24	48	24	nan	1.51	0.000	0.004	0.051
26	52	26	nan	1.49	0.000	0.004	0.051
28	56	28	nan	1.47	0.000	0.004	0.051
30	60	30	nan	1.45	0.000	0.004	0.051
40	80	40	nan	1.38	0.000	0.004	0.051
50	100	50	nan	1.33	0.000	0.004	0.050
60	120	60	nan	1.30	0.000	0.004	0.050
4	8	2	nan	5.57	0.000	0.001	0.000
6	12	3	nan	3.94	0.000	0.000	0.113
8	16	4	nan	3.29	0.000	0.000	0.083
10	20	5	nan	2.95	0.000	0.000	0.072
12	24	6	nan	2.72	0.000	0.000	0.065
14	28	7	nan	2.57	0.000	0.000	0.062
16	32	8	nan	2.45	0.000	0.000	0.060
18	36	9	nan	2.36	0.000	0.000	0.058
20	40	10	nan	2.29	0.000	0.000	0.057
22	44	11	nan	2.23	0.000	0.000	0.056
24	48	12	nan	2.17	0.000	0.000	0.055
26	52	13	nan	2.13	0.000	0.000	0.054
28	56	14	nan	2.09	0.000	0.000	0.054
30	60	15	nan	2.06	0.000	0.000	0.054
40	80	20	nan	1.93	0.000	0.001	0.052
50	100	25	nan	1.86	0.000	0.001	0.052
60	120	30	nan	1.80	0.000	0.001	0.052

Table 3: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 3$, and $\alpha = 0.05$

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	12	12	nan	1.51	0.000	0.016	0.060
6	18	18	nan	1.33	0.000	0.015	0.055
8	24	24	nan	1.23	0.000	0.014	0.053
10	30	30	nan	1.17	0.000	0.014	0.052
12	36	36	nan	1.13	0.000	0.014	0.052
14	42	42	nan	1.10	0.000	0.013	0.051
16	48	48	nan	1.07	0.000	0.013	0.051
18	54	54	nan	1.05	0.000	0.013	0.051
20	60	60	nan	1.03	0.000	0.013	0.051
22	66	66	nan	1.02	0.000	0.013	0.051
24	72	72	nan	1.00	0.000	0.013	0.051
26	78	78	nan	0.99	0.000	0.013	0.050
28	84	84	nan	0.98	0.000	0.013	0.050
30	90	90	nan	0.97	0.000	0.013	0.050
40	120	120	nan	0.94	0.000	0.013	0.050
50	150	150	nan	0.92	0.000	0.013	0.050
60	180	180	nan	0.90	0.000	0.012	0.050
4	12	8	nan	2.18	0.000	0.009	0.061
6	18	12	nan	1.88	0.000	0.009	0.055
8	24	16	nan	1.72	0.000	0.008	0.053
10	30	20	nan	1.63	0.000	0.008	0.052
12	36	24	nan	1.56	0.000	0.008	0.051
14	42	28	nan	1.51	0.000	0.008	0.050
16	48	32	nan	1.47	0.000	0.008	0.050
18	54	36	nan	1.44	0.000	0.008	0.050
20	60	40	nan	1.41	0.000	0.008	0.050
22	66	44	nan	1.39	0.000	0.008	0.050
24	72	48	nan	1.37	0.000	0.008	0.050
26	78	52	nan	1.35	0.000	0.007	0.050
28	84	56	nan	1.34	0.000	0.008	0.050
30	90	60	nan	1.33	0.000	0.008	0.050
40	120	80	nan	1.28	0.000	0.008	0.050
50	150	100	nan	1.24	0.000	0.007	0.050
60	180	120	nan	1.22	0.000	0.007	0.050

Table 3: Upper percentiles and type I errors when $p_1 : p_2 = 1 : 3$, and $\alpha = 0.05$ (continued)

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	12	4	nan	4.08	0.000	0.002	0.076
6	18	6	nan	3.30	0.000	0.002	0.060
8	24	8	nan	2.94	0.000	0.002	0.055
10	30	10	nan	2.72	0.000	0.002	0.052
12	36	12	nan	2.58	0.000	0.002	0.052
14	42	14	nan	2.47	0.000	0.002	0.051
16	48	16	nan	2.39	0.000	0.002	0.051
18	54	18	nan	2.32	0.000	0.002	0.050
20	60	20	nan	2.27	0.000	0.002	0.050
22	66	22	nan	2.22	0.000	0.002	0.050
24	72	24	nan	2.18	0.000	0.002	0.050
26	78	26	nan	2.15	0.000	0.002	0.049
28	84	28	nan	2.12	0.000	0.002	0.049
30	90	30	nan	2.10	0.000	0.002	0.049
40	120	40	nan	2.00	0.000	0.002	0.049
50	150	50	nan	1.94	0.000	0.002	0.049
60	180	60	nan	1.90	0.000	0.002	0.049
4	12	2	nan	7.95	0.000	0.000	0.000
6	18	3	nan	5.61	0.000	0.000	0.107
8	24	4	nan	4.72	0.000	0.000	0.077
10	30	5	nan	4.23	0.000	0.000	0.066
12	36	6	nan	3.93	0.000	0.000	0.062
14	42	7	nan	3.70	0.000	0.000	0.059
16	48	8	nan	3.54	0.000	0.000	0.057
18	54	9	nan	3.41	0.000	0.000	0.056
20	60	10	nan	3.31	0.000	0.000	0.055
22	66	11	nan	3.23	0.000	0.000	0.054
24	72	12	nan	3.15	0.000	0.000	0.053
26	78	13	nan	3.09	0.000	0.000	0.052
28	84	14	nan	3.04	0.000	0.000	0.052
30	90	15	nan	2.99	0.000	0.000	0.052
40	120	20	nan	2.82	0.000	0.000	0.051
50	150	25	nan	2.72	0.000	0.000	0.051
60	180	30	nan	2.64	0.000	0.000	0.050

Table 4: Upper percentiles and type I errors when $p_1 : p_2 = 2 : 1$, and $\alpha = 0.05$

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	2	12	1.71	0.42	0.050	0.058	0.083
6	3	18	1.29	0.35	0.050	0.049	0.068
8	4	24	1.08	0.31	0.050	0.045	0.062
10	5	30	0.97	0.29	0.050	0.044	0.059
12	6	36	0.89	0.27	0.050	0.042	0.057
14	7	42	0.83	0.26	0.050	0.041	0.055
16	8	48	0.78	0.25	0.050	0.041	0.054
18	9	54	0.75	0.24	0.050	0.041	0.054
20	10	60	0.72	0.24	0.050	0.041	0.054
22	11	66	0.70	0.23	0.050	0.040	0.053
24	12	72	0.68	0.22	0.050	0.040	0.052
26	13	78	0.66	0.22	0.050	0.040	0.052
28	14	84	0.64	0.22	0.050	0.039	0.052
30	15	90	0.63	0.21	0.050	0.039	0.052
40	20	120	0.58	0.20	0.050	0.039	0.051
50	25	150	0.55	0.19	0.050	0.039	0.051
60	30	180	0.53	0.18	0.050	0.038	0.050
4	2	8	18.95	0.59	0.050	0.058	0.101
6	3	12	9.27	0.49	0.050	0.046	0.077
8	4	16	6.38	0.43	0.050	0.041	0.066
10	5	20	5.05	0.40	0.050	0.039	0.063
12	6	24	4.30	0.37	0.050	0.037	0.059
14	7	28	3.80	0.35	0.050	0.037	0.057
16	8	32	3.45	0.34	0.050	0.036	0.056
18	9	36	3.17	0.33	0.050	0.036	0.055
20	10	40	2.98	0.32	0.050	0.036	0.055
22	11	44	2.81	0.31	0.050	0.035	0.054
24	12	48	2.68	0.30	0.050	0.035	0.054
26	13	52	2.58	0.30	0.050	0.035	0.054
28	14	56	2.48	0.29	0.050	0.035	0.054
30	15	60	2.40	0.29	0.050	0.034	0.053
40	20	80	2.13	0.27	0.050	0.034	0.052
50	25	100	1.96	0.26	0.050	0.033	0.051
60	30	120	1.84	0.25	0.050	0.033	0.051

Table 4: Upper percentiles and type I errors when $p_1 : p_2 = 2 : 1$, and $\alpha = 0.05$ (continued)

p_1	p_2	N	$U(\alpha)$	$U_H(\alpha)$	P_U	P_H	P_{BH}
4	2	4	nan	1.04	0.000	0.053	0.156
6	3	6	nan	0.82	0.000	0.035	0.103
8	4	8	nan	0.71	0.000	0.029	0.084
10	5	10	nan	0.64	0.000	0.026	0.075
12	6	12	nan	0.60	0.000	0.025	0.069
14	7	14	nan	0.56	0.000	0.024	0.065
16	8	16	nan	0.54	0.000	0.023	0.061
18	9	18	nan	0.52	0.000	0.023	0.060
20	10	20	nan	0.50	0.000	0.022	0.058
22	11	22	nan	0.49	0.000	0.022	0.058
24	12	24	nan	0.47	0.000	0.022	0.057
26	13	26	nan	0.46	0.000	0.022	0.056
28	14	28	nan	0.45	0.000	0.022	0.055
30	15	30	nan	0.44	0.000	0.021	0.055
40	20	40	nan	0.41	0.000	0.021	0.053
50	25	50	nan	0.39	0.000	0.021	0.053
60	30	60	nan	0.38	0.000	0.021	0.052
4	2	2	nan	1.83	0.000	0.046	0.000
6	3	3	nan	1.30	0.000	0.016	0.176
8	4	4	nan	1.08	0.000	0.011	0.125
10	5	5	nan	0.95	0.000	0.009	0.102
12	6	6	nan	0.87	0.000	0.009	0.089
14	7	7	nan	0.81	0.000	0.008	0.080
16	8	8	nan	0.77	0.000	0.008	0.075
18	9	9	nan	0.73	0.000	0.008	0.071
20	10	10	nan	0.71	0.000	0.008	0.069
22	11	11	nan	0.68	0.000	0.008	0.066
24	12	12	nan	0.66	0.000	0.007	0.064
26	13	13	nan	0.65	0.000	0.007	0.062
28	14	14	nan	0.63	0.000	0.007	0.061
30	15	15	nan	0.62	0.000	0.007	0.060
40	20	20	nan	0.57	0.000	0.007	0.057
50	25	25	nan	0.54	0.000	0.007	0.055
60	30	30	nan	0.52	0.000	0.007	0.054