Group Sequential Test for Checking Difference between Variances of Two Normal Populations

by

Tsunehisa IMADA*

(Received October 31, 2015)

Abstract

In this study we discuss a group sequential test for checking the difference between variances of two normal populations. Specifically, we construct repeated confidence boundaries for the group sequential test. Since it is difficult to determine them satisfying a specified significance level exactly, we determine conservative repeated confidence boundaries. We give simulation results regarding repeated confidence boundaries and power of the test.

key words: Conservation; Power of the test; Repeated confidence boundaries

1. Introduction

There are independent normal random variables X_1, X_2 satisfying

$$X_1 \sim N(\mu_1, \sigma_1^2),\ X_2 \sim N(\mu_2, \sigma_2^2).$$

We occasionally want to test the difference between σ_1^2 and σ_2^2. For example, when we test the difference between μ_1 and μ_2, the method for testing depends on whether $\sigma_1^2 = \sigma_2^2$ or not. On the other hand, assume two treatments are evaluated by X_1 and X_2 respectively. If $\mu_1 = \mu_2$, the treatment having smaller variance is preferable, because it indicates the uniform effect for various cases.

Group sequential procedure is a method to test the difference between two population parameters by using grouped observations sequentially stage by stage. If the remarkable difference appears at early stage, we can terminate the test with small number of observations. Group sequential procedure is characterized by its repeated confidence boundaries. Pocock (1977) [3] used constant repeated confidence boundaries through all stages. Lan and DeMets (1983) [2] constructed the flexible repeated confidence boundaries by allocating a specified significance level α at each stage by using α-error spending function. Various studies on group sequential procedures are introduced in detail by Jennison and Turnbull (1999) [1].

The aim of this study is to construct group sequential procedures for checking the difference between σ_1^2 and σ_2^2. We consider two kinds of hypotheses for testing as follows.

$$H_0 : \sigma_1^2 = \sigma_2^2 \ \text{ vs. } \ H_1 : \sigma_1^2 \neq \sigma_2^2. \qquad (1.1)$$

$$H_0 : \sigma_1^2 = \sigma_2^2 \ \text{ vs. } \ H_1 : \sigma_1^2 > \sigma_2^2. \qquad (1.2)$$

(1.1) is the two-sided test and (1.2) is the one-sided test. We construct group sequential procedures for (1.1) and (1.2) respectively. However, it is difficult to determine repeated confidence boundaries satisfying a specified significance level exactly. We determine conservative repeated confidence boundaries.

In Section 2, we discuss how to construct group sequential procedures for (1.1) and (1.2) respectively. In Section 3, we give simulation results regarding repeated confidence boundaries and power of the test. In Section 4, we give some concluding remarks.

2. Group sequential procedure

We assume that a group of data with sample size g_1 from $N(\mu_1, \sigma_1^2)$ and a group of data with sample size g_2 from $N(\mu_2, \sigma_2^2)$ are obtained at each stage for total number of tests K. Here g_1 may not be necessarily equal to g_2. Then the sample sizes obtained till the kth stage from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ are $n_{1,k} = kg_1$ and $n_{2,k} = kg_2$ respectively. Let

$$x_{1,n_{1,k-1}+1}, x_{1,n_{1,k-1}+2}, \cdots, x_{1,n_{1,k}}$$

and

$$x_{2,n_{2,k-1}+1}, x_{2,n_{2,k-1}+2}, \cdots, x_{2,n_{2,k}}$$

be samples at the kth stage from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Let

$$\bar{x}_{1,k} = \frac{\sum_{i=1}^{n_{1,k}} x_{1,i}}{n_{1,k}},\ \bar{x}_{2,k} = \frac{\sum_{i=1}^{n_{2,k}} x_{2,i}}{n_{2,k}}$$

and

$$\nu_{1,k}^2 = \frac{\sum_{i=1}^{n_{1,k}} (x_{1,i} - \bar{x}_{1,k})^2}{n_{1,k} - 1},$$

$$\nu_{2,k}^2 = \frac{\sum_{i=1}^{n_{2,k}} (x_{2,i} - \bar{x}_{2,k})^2}{n_{2,k} - 1}.$$

* Liberal Arts Education Center, Kumamoto Campus, Tokai University, Professor

Here we use the statistic

$$F_k = \frac{\nu_{1,k}^2}{\nu_{2,k}^2}$$

for testing at the kth stage. F_k is distributed according to F-distribution with degrees of freedom $(n_{1,k}-1, n_{2,k}-1)$ under H_0.

For the hypotheses (1.1) we specify the repeated confidence boundaries $b_{1,k}, b_{2,k}$ satisfying $0 < b_{1,k} < b_{2,k}$ $(k = 1, 2, \ldots, K)$ and carry out the group sequential test as follows:
(1) At the k $(1 \le k \le K-1)$th stage
Case 1. If $F_k < b_{1,k}$ or $b_{2,k} < F_k$, we stop the test with rejection of H_0.
Case 2. If $b_{1,k} \le F_k \le b_{2,k}$, we go to the $(k+1)$th stage.
(2) At the final Kth stage
Case 1. If $F_K < b_{1,K}$ or $b_{2,K} < F_K$, we reject H_0.
Case 2. If $b_{1,K} \le F_K \le b_{2,K}$, we do not reject H_0.

For the hypotheses (1.2) we specify the repeated confidence boundaries $b_k (> 0)$ $(k = 1, 2, \ldots, K)$ and carry out the group sequential test as follows:
(1) At the k $(1 \le k \le K-1)$th stage
Case 1. If $F_k > b_k$, we stop the test with rejection of H_0.
Case 2. If $F_k \le b_k$, we go to the $(k+1)$th stage.
(2) At the final Kth stage
Case 1. If $F_K > b_K$, we reject H_0.
Case 2. If $F_K \le b_K$, we do not reject H_0.

Next, we discuss how to determine the repeated confidence boundaries for a specified significance level. For the hypotheses (1.1) the probability that H_0 is rejected is

$$\begin{aligned} & P(F_1 < b_{1,1} \text{ or } b_{2,1} < F_1) \\ & + P(b_{1,1} \le F_1 \le b_{2,1},\ F_2 < b_{1,2} \text{ or } b_{2,2} < F_2) \\ & + P(b_{1,1} \le F_1 \le b_{2,1},\ b_{1,2} \le F_2 \le b_{2,2}, \\ & \quad F_3 < b_{1,3} \text{ or } b_{2,3} < F_3) + \cdots\cdots \\ & + P(b_{1,1} \le F_1 \le b_{2,1},\ \ldots,\ b_{1,K-1} \le F_{K-1} \le b_{2,K-1}, \\ & \quad F_K < b_{1,K} \text{ or } b_{2,K} < F_K). \end{aligned} \tag{2.1}$$

Although we should determine $b_{1,k}, b_{2,k}$ $(k = 1, 2, \ldots, K)$ so that (2.1) may be equal to a specified significance level α under H_0, it is difficult to formulate (2.1). However, we can determine conservative repeated confidence boundaries easily.

$$\sum_{k=1}^{K} P(F_k < b_{1,k} \text{ or } b_{2,k} < F_k) \tag{2.2}$$

is larger than (2.1), because

$$\begin{aligned} & P(b_{1,1} \le F_1 \le b_{2,1},\ F_2 < b_{1,2} \text{ or } b_{2,2} < F_2) \\ & \le P(F_2 < b_{1,2} \text{ or } b_{2,2} < F_2), \\ & P(b_{1,1} \le F_1 \le b_{2,1},\ b_{1,2} \le F_2 \le b_{2,2}, \\ & \quad F_3 < b_{1,3} \text{ or } b_{2,3} < F_3) \\ & \le P(F_3 < b_{1,3} \text{ or } b_{2,3} < F_3), \\ & \vdots \\ & P(b_{1,1} \le F_1 \le b_{2,1},\ \ldots,\ b_{1,K-1} \le F_{K-1} \le b_{2,K-1}, \\ & \quad F_K < b_{1,K} \text{ or } b_{2,K} < F_K) \\ & \le P(F_K < b_{1,K} \text{ or } b_{2,K} < F_K). \end{aligned}$$

If we determine $b_{1,k}, b_{2,k}$ so that

$$P(F_k < b_{1,k} \text{ or } b_{2,k} < F_k) = \frac{\alpha}{K} \tag{2.3}$$

for $k = 1, 2, \ldots, K$, (2.1) is smaller than α under H_0. Specifically, $b_{1,k}, b_{2,k}$ $(k = 1, 2, \ldots, K)$ are conservative repeated confidence boundaries. If we determine $b_{1,k}, b_{2,k}$ $(k = 1, 2, \ldots, K)$ so that

$$P(F_k < b_{1,k}) = \frac{\alpha}{2K},\ P(b_{2,k} < F_k) = \frac{\alpha}{2K} \tag{2.4}$$

using F-distribution with degrees of freedom $(n_{1,k} - 1, n_{2,k} - 1)$, the condition (2.3) is satisfied. If $g_1 = g_2$, by (2.4)

$$b_{1,k} = \frac{1}{b_{2,k}}.$$

For the hypotheses (1.2), if we determine b_k so that

$$P(b_k < F_k) = \frac{\alpha}{K} \tag{2.5}$$

for $k = 1, 2, \ldots, K$, b_ks are conservative repeated confidence boundaries.

3. Simulation results

In this section we give some numerical examples regarding the repeated confidence boundaries for a specified significance level and the power of the test. Let $\alpha = 0.05$ and $K = 4, 5$. Assume $g_1 = g_2 = g = 10, 20$. Tables 1,2 give the repeated confidence boundaries $b_{2,k}$s for (1.1) and $K = 4, 5$. ($b_{1,k}$ is determined by $b_{1,k} = 1/b_{2,k}$.) Tables 3,4 give the repeated confidence boundaries b_ks for (1.2) and $K = 4, 5$.

Table 1: Repeated confidence boundaries $b_{2,k}$s for the hypotheses (1.1) and $K = 4$

Stage k	1	2	3	4
$g = 10$	6.139	3.299	2.593	2.262
$g = 20$	3.299	2.262	1.933	1.764

Table 2: Repeated confidence boundaries $b_{2,k}$s for the hypotheses (1.1) and $K = 5$

Stage k	1	2	3	4	5
$g = 10$	6.542	3.432	2.674	2.322	2.114
$g = 20$	3.432	2.322	1.974	1.796	1.686

Table 3: Repeated confidence boundaries b_ks for the hypotheses (1.2) and $K = 4$

Stage k	1	2	3	4
$g = 10$	5.005	2.903	2.345	2.076
$g = 20$	2.903	2.076	1.805	1.664

Table 4: Repeated confidence boundaries b_ks for the hypotheses (1.2) and $K = 5$

Stage k	1	2	3	4	5
$g = 10$	5.352	3.028	2.424	2.136	1.963
$g = 20$	3.028	2.136	1.846	1.696	1.602

Since the repeated confidence boundaries in Tables 1 to 4 are conservative, we calculate the actual Type I error by Monte Carlo simulation. Table 5 gives Type I error calculated by Monte Carlo simulation with 1,000,000 times of experiments. Type I error for $K = 5$ is smaller compared to that for $K = 4$ in each case. Specifically, the repeated confidence boundaries for $K = 5$ are more conservative compared to them for $K = 4$. Tables 6 to 9 give the probability that H_0 is rejected at kth stage for (1.1) and (1.2) respectively. It decreases as k increases from 1 to K.

Table 5: Type I error

Hypotheses	(1.1)		(1.2)	
K	4	5	4	5
$g = 10$	0.0367	0.0340	0.0346	0.0325
$g = 20$	0.0360	0.0331	0.0342	0.0318

Table 6: Probability that H_0 is rejected at kth stage for the hypotheses (1.1) and $K = 4$

stage k	1	2	3	4
$g = 10$	0.0124	0.0103	0.0078	0.0063
$g = 20$	0.0125	0.0098	0.0075	0.0062

Table 7: Probability that H_0 is rejected at kth stage for the hypotheses (1.1) and $K = 5$

stage k	1	2	3	4	5
$g = 10$	0.0098	0.0082	0.0065	0.0052	0.0043
$g = 20$	0.0098	0.0078	0.0062	0.0051	0.0042

Table 8: Probability that H_0 is rejected at kth stage for the hypotheses (1.2) and $K = 4$

stage k	1	2	3	4
$g = 10$	0.0125	0.0095	0.0071	0.0056
$g = 20$	0.0123	0.0094	0.0070	0.0055

Table 9: Probability that H_0 is rejected at kth stage for the hypotheses (1.2) and $K = 5$

stage k	1	2	3	4	5
$g = 10$	0.0101	0.0079	0.0058	0.0048	0.0039
$g = 20$	0.0101	0.0075	0.0058	0.0046	0.0038

Next, we consider the power of the test. To calculate the power we should specify the difference between σ_1^2 and σ_2^2 under H_1. Let

$$\sigma_2^2 = \gamma\sigma_1^2 \text{ with } \gamma = 0.9, 0.8, 0.7, 0.6, 0.5$$

under H_1. Since we obtain the power using Monte Carlo simulation, we specify the values of μ_1, μ_2 and σ_1^2 respectively as

$$\mu_1 = 0,\ \mu_2 = 0 \text{ and } \sigma_1^2 = 1.$$

Tables 10,11 give the power of the test for (1.1) and $K = 4, 5$. Tables 12,13 give the power of the test for (1.2) and $K = 4, 5$. We continue 100,000 times of experiments to obtain the power by Monte Carlo simulation. The power increases as γ decreases from 0.9 to 0.5. Specifically, the power increases as the ratio σ_2^2/σ_1^2 decreases. The power using $g = 20$ is uniformly higher compared to that using $g = 10$.

Table 10: Power of the test for the hypotheses (1.1) and $K = 4$

γ	0.9	0.8	0.7	0.6	0.5
$g = 10$	0.063	0.176	0.433	0.770	0.967
$g = 20$	0.096	0.353	0.776	0.981	1.000

Table 11: Power of the test for the hypotheses (1.1) and $K = 5$

γ	0.9	0.8	0.7	0.6	0.5
$g = 10$	0.066	0.203	0.517	0.857	0.989
$g = 20$	0.106	0.418	0.857	0.994	1.000

Table 12: Power of the test for the hypotheses (1.2) and $K = 4$

γ	0.9	0.8	0.7	0.6	0.5
$g = 10$	0.097	0.255	0.543	0.847	0.983
$g = 20$	0.146	0.463	0.851	0.991	1.000

Table 13: Power of the test for the hypotheses (1.2) and $K = 5$

γ	0.9	0.8	0.7	0.6	0.5
$g = 10$	0.099	0.290	0.621	0.907	0.995
$g = 20$	0.159	0.527	0.908	0.998	1.000

4. Conclusions

In this study we have discussed the group sequential procedures for checking the difference between variances of two normal populations. We determined the conservative repeated confidence boundaries using certain inequalities regarding probabilities. We gave some simulation results regarding the repeated confidence boundaries and the power of the test and investigated their characteristics. We confirmed that the repeated confidence boundaries are fairly conservative. We should develop less conservative repeated confidence boundaries using various devices in the future.

Acknowledgement

The author is deeply grateful to the referee and the editors for their valuable comments and suggestions.

References

[1] C. Jennison and B. W. Turnbull, "Group sequential methods with applications to clinical trials." Chapman & Hall/CRC Interdisciplinary Statistics, (1999).
[2] K. K. G. Lan and D. L. DeMets, *Discrete sequential boundaries for clinical trials*, Biometrika **70** (1983), 659-663.
[3] S. J. Pocock, *Group sequential methods in the design and analysis of clinical trials*, Biometrika **64** (1977), 191-199.