The Korean Journal of Applied Statistics (2010)
23(2), 317–323

Unbalanced ANOVA for Testing Shape Variability in Statistical Shape Analysis

Jong Geon Kim[1] · Yong-Seok Choi[2] · Nae-Young Lee[3]

[1]School of Computer and Information, Busan College of Information Technology
[2]Department of Statistics, Pusan National University; [3]Department of Nursing, Silla University

(Received March 2010; accepted March 2010)

Abstract

Measures are very useful tools for comparing the shape variability in statistical shape analysis. For examples, the Procrustes statistic(PS) is isolated measure, and the mean Procrustes statistic(MPS) and the root mean square measure(RMS) are overall measures. But these measures are very subjective, complicated and moreover these measures are not statistical for comparing the shape variability. Therefore we need to study some tests. It is well known that the Hotelling's T^2 test is used for testing shape variability of two independent samples. And for testing shape variabilities of several independent samples, instead of the Hotelling's T^2 test, one way analysis of variance(ANOVA) can be applied. In fact, this one way ANOVA is based on the balanced samples of equal size which is called as BANOVA. However, If we have unbalanced samples with unequal size, we can not use BANOVA. Therefore we propose the unbalanced analysis of variance(UNBANOVA) for testing shape variabilities of several independent samples of unequal size.

Keywords: Hotelling's T^2 test, mean shape, shape variability, UNBANOVA.

1. Introduction

The shape is all the geometrical information that remains when location, size and rotational effects are filtering out from an object (Kendall, 1984). That is, two objects have the same shape if they can be translated, scaled and rotated to each other such that they match exactly. In a wide variety of disciplines, it is of greate practical importance to measure, describe and compare the shapes of objects. Procrustes analysis is a very useful tool for doing these and estimating a mean shape.

When several objects are fitted using Procrustes analysis the method has been called generalized Procrustes analysis(GPA) (Gower, 1975), whereas when a single objects is fitted to one other, the method has been called ordinary Procrustes analysis(OPA).

Procrustes analysis involves fitting configurations with similarity transformations to be as close as possible according to Euclidean distance, and using least square(LS) technique. So, Choi *et al.*

This work was supported by the Korea Research Foundation Grant funded by the Korean Government(MOEHRD, Basic Research Promotion Fund)(KRF-2007-313-C00117).
[2]Corresponding author: Professor, Department of Statistics, Pusan National University, Jangjeon-Dong, Geumjeong-Gu, Busan 609-735, Korea. E-mail: yschoi@pusan.ac.kr

(2005) provided the Procrustes statistic for comparing configurations. Hyun (2006) provided the methods by visualizing the mean shapes. And Kim (2009) also provided the isolated and overall measures. Of course, with respect to objective quantitative, the comparison of the shape variability using these measures is better than using the naked eyes. But we noted that these measures also were very subjective, complicated and these measures were not statistical for comparing the shape variability. Therefore in this article, we study some tests. Section 2 provides an introduction of Procrustes analysis. In Section3, we introduce some tests for shape variability and we propose an unbalanced ANOVA(UNBANOVA) for unbalanced samples of unequal size. Section 4 gives two examples. Finally, concluding remark is given in Section 5.

2. Procrustes Analysis

Procrustes analysis is very useful tools for analyzing landmark data and it involves fitting the objects using similarity transformations to be as close as possible to one another usually by minimizing the sum of squared Euclidean distances. In general, there are two techniques in Procrustes analysis. One is OPA, which is used for fitting and comparing two configurations. Another is GPA for obtaining a mean shape and to explore the structure of shape variability in a data set of more than two configurations.

Firstly, for the OPA consider two centred $k \times m$ matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ of coordinates from k landmarks in m dimensions. The estimation of α, $\mathbf{\Gamma}$ and β by OPA method based on least squares(LS OPA) is carried out by minimizing the squared Euclidean distance

$$O(\mathbf{X}_1, \mathbf{X}_2) = \left\| \mathbf{X}_2 - \beta \mathbf{X}_1 \mathbf{\Gamma} - \mathbf{1}_k \alpha^T \right\|^2,$$

where $\|\mathbf{A}\| = \sqrt{\mathrm{tr}(\mathbf{A}^T \mathbf{A})}$ is the Euclidean norm, α is an $m \times 1$ location vector, $\mathbf{\Gamma}$ is an $m \times m$ rotation matrix such that $|\mathbf{\Gamma}| = 1$ and $\beta > 0$ is a scale parameter (Dryden and Mardia, 1998).

Secondly, for the GPA method by least squares(LS GPA) with $n \geq 2$ configurations of k landmarks in $m \geq 2$ dimensions, let $k \times m$ configurations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be centred and scaled. We wish to minimize following objective function,

$$G(\mathbf{X}_1, \ldots, \mathbf{X}_n) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left\| \left(\beta_i \mathbf{X}_i \mathbf{\Gamma}_i + \mathbf{1}_k \alpha_i^T \right) - \left(\beta_j \mathbf{X}_j \mathbf{\Gamma}_j + \mathbf{1}_k \alpha_j^T \right) \right\|^2, \tag{2.1}$$

subject to a constraint on the size of the average, $C(\bar{\mathbf{X}}) = 1$, where α_i(location parameter), $\mathbf{\Gamma}_i$(rotation matrix such that $|\mathbf{\Gamma}| = 1$), $\beta_i > 0$(scale parameter), $i = 1, \ldots, n$, $\|\mathbf{X}\| = \sqrt{\mathrm{tr}(\mathbf{X}^T \mathbf{X})}$ and $C(\mathbf{X})$ is the centroid size and the average configuration is $\bar{\mathbf{X}} = 1/n \sum_{i=1}^{n} \left(\beta_i \mathbf{X}_i \mathbf{\Gamma}_i + \mathbf{1}_k \alpha_i^T \right)$. Note that

$$\begin{aligned} G(\mathbf{X}_1, \ldots, \mathbf{X}_n) &= \inf_{(\beta_i, \mathbf{\Gamma}_i, \alpha_i)} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left\| \left(\beta_i \mathbf{X}_i \mathbf{\Gamma}_i + \mathbf{1}_k \alpha_i^T \right) - \left(\beta_j \mathbf{X}_j \mathbf{\Gamma}_j + \mathbf{1}_k \alpha_j^T \right) \right\|^2 \\ &= \inf_{(\beta_i, \mathbf{\Gamma}_i, \alpha_i)} \sum_{i=1}^{n} \left\| \left(\beta_i \mathbf{X}_i \mathbf{\Gamma}_i + \mathbf{1}_k \alpha_i^T \right) - \frac{1}{n} \sum_{j=1}^{n} \left(\beta_j \mathbf{X}_j \mathbf{\Gamma}_j + \mathbf{1}_k \alpha_j^T \right) \right\|^2 . \end{aligned}$$

In general, the estimate of mean shape, denoted by $\hat{\mu}$, can be defined as the average of configurations after the objects have been translated, rotated and scaled. Thus optional solution to minimization of Equation (2.1) over μ is given by

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_i \mathbf{X}_i \hat{\mathbf{\Gamma}}_i + \mathbf{1}_k \hat{\alpha}_i^T\right).$$

An explicit eigenvector solution to the mean shape for two dimensional data can be found as the complex eigenvector corresponding to the largest eigenvalue of the complex sum of squares and products matrix (Kent, 1994). The Procrustes mean shape μ has to be found iteratively for $m \geq 3$ dimensional data. For a practical implementation when the objects are in $m > 2$ dimensions one could use the GPA algorithm of Gower (1975), which is modified by Ten Berge (1977).

3. Tests for Shape Variability

3.1. Two independent samples

Consider two independent random samples $\mathbf{X}_{j1}, \ldots, \mathbf{X}_{jn_j}$, $j = 1, 2$ (each a $k \times m$ matrix) are taken from an isotropic normal model with mean μ_j and transformed by an additional location, rotation and scale, *i.e.*

$$\mathbf{X}_{ji} = \beta_{ji}\left(\mu_j + \mathbf{E}_{ji}\right)\mathbf{\Gamma}_{ji} + \mathbf{1}_k \alpha_{ji}^T, \qquad \text{vec}\left(\mathbf{E}_{ji}\right) \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}_{km}\right), \tag{3.1}$$

where $\beta_{ji} \geq 0$ (scale), $\mathbf{\Gamma}_{ji} \in SO(m)$ with $SO(m)$ is the special orthogonal group of rotations, $\alpha_{ji} \in \mathbb{R}^m$(translation). Both populations are assumed to have a common variance for each coordinate σ^2. To test

$$H_0 : \mu_1 = \mu_2 \quad \text{versus} \quad H_1 : \mu_1 \neq \mu_2$$

we could carry out the Hotelling's T^2 two sample test in the shape space. Let $\mathbf{v}_i$, $i = 1, \ldots, n_1$ and $\mathbf{w}_j$, $j = 1, \ldots, n_2$ be the complex configurations with mean shape $\hat{\mu}$. The multivariate normal model is proposed in the shape space, where

$$\mathbf{v}_i \sim N\left(\delta_1, \mathbf{\Sigma}\right), \qquad \mathbf{w}_j \sim N\left(\delta_2, \mathbf{\Sigma}\right), \quad i = 1, \ldots, n_1;\ j = 1, \ldots, n_2$$

and the $\mathbf{v}_i$ and $\mathbf{w}_j$ are all mutually independent, and common covariance matrices are assumed. Since the dimension of the shape space is $M = km - m - m(m-1)/2 - 1$ and the length of each vector $\mathbf{v}_i$ and $\mathbf{w}_j$ is $(k-1)m \geq M$, we have singular covariance matrices. We write $\bar{\mathbf{v}}$, $\bar{\mathbf{w}}$ and $\mathbf{S}_\mathbf{v}$, $\mathbf{S}_\mathbf{w}$ for the sample means and sample covariance matrices (with divisors n_1 and n_2) in each group. The Mahalanobis distance squared between $\bar{\mathbf{v}}$ and $\bar{\mathbf{w}}$ is

$$D^2 = \left(\bar{\mathbf{v}} - \bar{\mathbf{w}}\right)^T \mathbf{S}_\mathbf{u}^- \left(\bar{\mathbf{v}} - \bar{\mathbf{w}}\right),$$

where $\mathbf{S}_\mathbf{u} = \left(n_1 \mathbf{S}_\mathbf{v} + n_2 \mathbf{S}_\mathbf{w}\right) / \left(n_1 + n_2 - 2\right)$, and $\mathbf{S}_\mathbf{u}^-$ is the Moore-Penrose generalized inverse of $\mathbf{S}_\mathbf{u}$. The rank of $\mathbf{S}_\mathbf{u}$ is $\min(M, n_1 - 1, n_2 - 1)$ and we assume that the rank of our sample covariance matrices is M. Under H_0 we have $\delta_1 = \delta_2$, and we use the test statistic

$$F = \frac{n_1 n_2 \left(n_1 + n_2 - M - 1\right)}{\left(n_1 + n_2\right)\left(n_1 + n_2 - 2\right) M} D^2.$$

The test statistic has an $F_{M, n_1 + n_2 - M - 1}$ distribution under H_0. Hence, we reject H_0 for large values of F.

3.2. Several independent samples

Goodall (1991) suggested a balanced analysis of variance with independent random samples $\mathbf{X}_{j1},\ldots,\mathbf{X}_{jn}$, $j=1,\ldots,n_G$ each of size n from n_G populations defined in Equation (3.1) with mean μ_j. Let $\hat{\mu}_j$ be the LS mean shape of the jth group by LS GPA and $\hat{\mu}$ is the overall pooled Procrustes mean shape. A suitable test statistic is

$$F=\frac{n(n-1)n_G\sum_{j=1}^{n_G}d_F^2(\hat{\mu}_j,\hat{\mu})}{(n_G-1)\sum_{j=1}^{n_G}\sum_{i=1}^{n}d_F^2(\mathbf{X}_{ji},\hat{\mu}_j)}.$$

Here, $d_F^2(\mathbf{v},\mathbf{w})$ is the full Procrustes distance between complex configurations $\mathbf{v}$ and $\mathbf{w}$ given by

$$d_F^2(\mathbf{v},\mathbf{w})=\left[1-\frac{\mathbf{w}^*\mathbf{v}\mathbf{v}^*\mathbf{w}}{\mathbf{v}^*\mathbf{v}\mathbf{w}^*\mathbf{w}}\right]^{\frac{1}{2}},$$

where $\mathbf{v}^*$ and $\mathbf{w}^*$ denote transpose of the complex conjugates of $\mathbf{v}$ and $\mathbf{w}$, respectively. Therefore $\sum_{i=1}^{n}\sum_{j=1}^{n_G}d_F^2(\mathbf{X}_{ji},\hat{\mu}_j)$ has an approximate chi-squared distribution with $(n_G-1)M$ degree of freedom and $n\sum_{j=1}^{n_G}d_F^2(\hat{\mu}_j,\hat{\mu})$ has an approximate chi-square distribution with $(n-1)n_GM$ degree of freedom. So the test statistic has the approximate F-distribution under the null hypothesis of equal means with $((n_G-1)M,(n-1)n_GM)$ degrees of freedom and the null hypothesis is rejected for large values of the statistic.

3.3. Unbalanced ANOVA

In Section 3.2, we described the BANOVA for testing mean shapes of several independent random samples of equal size. If we have an unbalanced several independent random samples of unequal size, then we can't use the BANOVA any longer. Therefor we propose a suitable ANOVA.

For this, consider an unbalanced analysis of variance with independent random samples $\mathbf{X}_{j1},\ldots,\mathbf{X}_{jn_j}$, each of size n_j, $j=1,\ldots,n_G$ from n_G populations defined in Equation (3.1) with mean μ_j. Then we can calculate the total sum of squares

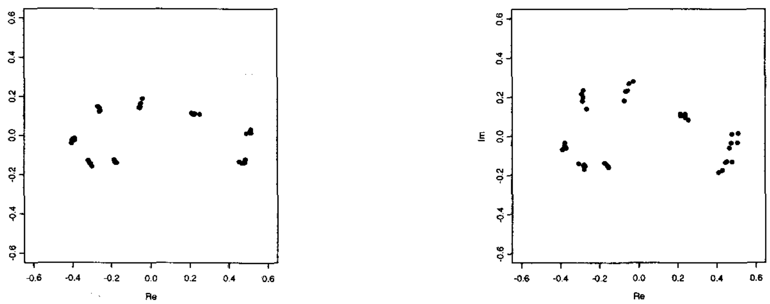
$$\mathrm{TSS}=\sum_{i=1}^{n_j}\sum_{j=1}^{n_G}d_F^2(\mathbf{X}_{ji},\hat{\mu}),$$

which has an approximate chi-square distribution with $(N-1)M$ degree of freedom. Here $N=\sum_{j=1}^{n_G}n_j$ is the number of all configurations. And also we have the sum of squares of treatment

$$\mathrm{SST}=\sum_{i=1}^{n_j}\sum_{j=1}^{n_G}d_F^2(\hat{\mu}_j,\hat{\mu}),$$

which has an approximate chi-square distribution with $(n_G-1)M$ degree of freedom. Therefore the sum of squares for error is

$$\mathrm{SSE}=\sum_{i=1}^{n_j}\sum_{j=1}^{n_G}d_F^2(\mathbf{X}_{ji},\hat{\mu}_j),$$



(a) (b)

Figure 4.1. The pre-shapes of the neural skulls of 10 rats: (a) The pre-shapes of the control group, (b) The pre-shapes of the hydrocephaly group.

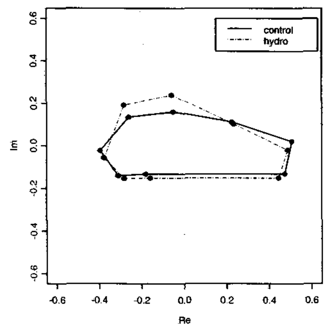


Figure 4.2. The overlap of the control group and hydrocephaly group LS mean shapes by LS GPA.

which has an approximate chi-square distribution with $(N - n_G)M$ degree of freedom. Thus we propose the test statistic for an unbalanced ANOVA

$$F = \frac{\mathrm{SST}/\,(n_G - 1)\,M}{\mathrm{SSE}/\,(N - n_G)\,M}.$$

Under the null hypothesis of equal means, the test statistic has an approximate F distribution with $((n_G - 1)M, (N - n_G)M)$ degrees of freedom and the null hypothesis is rejected for large values of the statistic. We call this unbalanced analysis of variance(UNBANOVA) from now on.

4. Examples

EXAMPLE 4.1. Consider the neural skulls of rats data data (Goodall, 1991). The data set is divided into a control group of five normal rats and a group of five hydrocephaly rats, with eight landmarks in two dimensions.

In Figure 4.1 we see the pre-shapes of the control and hydrocephaly group for comparing two configurations. We have the overlap of the control and hydrocephaly group LS mean shapes by LS GPA in Figure 4.2. This figure shows that the LS mean shape of the hydrocephaly group is larger relatively than the control group, in particular, a part of the front head at LS mean shape is large. We note that an approach using the Procrustes statistic is complicated and isolated. So in this

Table 4.1. The rates of rejecting $H_0(\%)$

$n_1 = n_2$	5	6	7	8	9	10	11	12	13
Rates	0	0	17	64	91	100	100	100	100

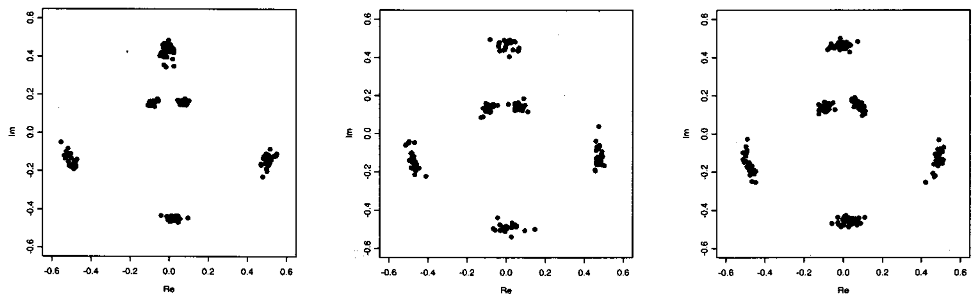


Figure 4.3. The pre-shapes of the mouse vertebra data: (a) The pre-shapes of the small group, (b) The pre-shapes of the large group, (c) The pre-shapes of the control group.

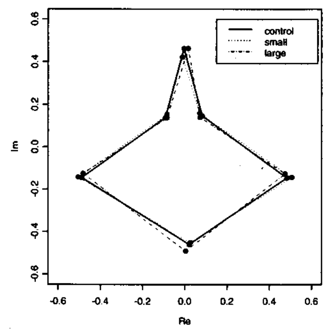


Figure 4.4. The overlap of the control group, small group and large group LS mean shapes by LS GPA.

example, we will consider the Hotelling's T^2 test for LS mean shapes. We have very strong evidence that the LS mean shapes $\hat{\mu}_C$ and $\hat{\mu}_H$ of two groups are equal because the observed test statistic is $F = 6.001$ and $P(F_{4,5} > 6.001) = 0.3062$. However through the figures, we can not guarantee the equality of their shapes obtaining from the result of the Hotelling's T^2 tests. In general, we must assume the multivariate normality of samples in the Hotelling's T^2 test. However this example doesn't satisfy this assumption because both sizes of two samples are very small. In addition, in calculating the test statistics F, the rank of $\mathbf{S_u}$ is assumed as $\min(M, n_1 - 1, n_2 - 1)$. In fact, since $M = 12$ and $n_1 = n_2 = 5$ in this example, the rank of $\mathbf{S_u}$ doesn't satisfy this assumption. Therefore to satisfy this assumption of rank, we need to make the new data generated by $(1-\varepsilon)N(\mu, \mathbf{\Sigma}_1) + \varepsilon N(\mu, \mathbf{\Sigma}_2)$, where $\mathbf{\Sigma}_1 = 0.2\mathbf{I}_2$, $\mathbf{\Sigma}_2 = 1.0\mathbf{I}_2$, $\varepsilon = 0.1$. Consider the Hotelling's T^2 test to add the new data. In Table 4.1, if $n_1 = n_2 = 13$, that is, the rank of $\mathbf{S_u}$ is 12, then we can guarantee the difference of two LS mean shapes via the Hotelling's T^2 test with 100% rate of rejecting $H_0 : \hat{\mu}_C = \hat{\mu}_H$.

EXAMPLE 4.2. The mouse vertebra data (Mardia and Dryden, 1989) are used in this example. Consider the mouse vertebra data. There are $n_1 = 30$ control group, $n_2 = 23$ small group and

$n_3 = 23$ large group with six landmarks in two dimensions. Figure 4.3 gives the pre-shapes of the the control group, the small group and large group by LS GPA. We have the overlap of the control group, the small group and large group LS mean shapes by LS GPA in Figure 4.4.

In this case, it is impossible to use the Hotelling's T^2 test for more than two LS mean shapes. Moreover, since $n_1 = 30$, $n_2 = 23$ and $n_3 = 23$, we use an UNBANOVA model proposed in Section 3.3 instead of a BANOVA model. In UNBANOVA model, since we have the observed test statistic $F = 0.2126$ and $P(F_{16,584} \geq 0.2126) = 0.0126$, we have very strong evidence that the LS mean shapes of three groups are different.

5. Concluding Remark

We have studied some tests in statistical shape analysis. For two independent samples, we have used the Hotelling's T^2 test. In general, we must assume the multivariate normality of samples in the Hotelling's T^2 test. However we met an example such that both sizes of two samples are very small and they did not satisfy the assumption of multivariate normality. Moreover, in calculating the test statistics F, we have a case that the assumption of the rank of $\mathbf{S_u}$ is not satisfied. In fact, we had these problems in the neural skull of 10 rats data. For overcoming this problem, we showed a simulation such that if $n_1 = n_2 = 13$, then we could guarantee the difference of two LS mean shapes via the Hotelling's T^2 test with 100% rate of rejecting $H_0 : \mu_C = \mu_H$.

Moreover, we noted that it is impossible to use the Hotelling's T^2 test for more than two LS mean shapes. Also, for several independent samples, instead of the Hotelling's T^2 test, one way ANOVA was applied for testing LS mean shapes $\hat{\mu}_1, \ldots, \hat{\mu}_{n_G}$. If we have unbalanced samples with unequal size, we could not use the BANOVA. So we have proposed the UNBANOVA.

References

Choi, Y. S., Hyun, G. H. and Yun, W. J. (2005). Biplots variability based on the Procrustes analysis, *Journal of the Korean Data Analysis Society*, **7**, 1925–1933.

Dryden, I. L. and Mardia, K. V. (1998). *Statistical Shape Analysis*, John Wiley & Sons.

Goodall, C. (1991). Prucrustes methods in the statistical analysis of shape(with discussion), *Journal of Royal Statistical Society, Series B*, **53**, 285–339.

Gower, J. C. (1975). Generalized Procrustes analysis, *Psychometrika*, **40**, 33–50.

Hyun, G. H. (2006). Resistant generalized Procrustes methods in shape analysis, *Ph.D thesis, Pusan National University*.

Kendall, D. G. (1984). Procrustean metrics, and complex projective spaces, *Bulletin of the London Mathematical Society*, **16**, 81–121.

Kent, J. T. (1994). The complex Bingham distribution and shape analysis, *Journal of Royal Statistical Society, Series B*, **56**, 285–299.

Kim, J. G. (2009). A study on measures and tests for shape variability in statistical shape analysis, *Ph.D thesis, Pusan National University*.

Mardia, K. V. and Dryden, I. L. (1989). The statistical analysis of shape data, *Biometrika*, **76**, 271–281.

Ten Berge, J. M. F. (1977). Orthogonal Procrustes rotation for two or more matrices, *Psychometrika*, **42**, 267–276.