S Afr Optom 2006 **65** (2) 62 – 67

Application of multivariate analysis of variance (MANOVA) to distance refractive variability and mean distance refractive state

S Abelman*

Centre for Numerical Analysis and Computational Mathematics, School of Computational and Applied Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa

< abelman@cam.wits.ac.za >

Abstract

Refractive state can be regarded as a dynamic quantity. Multiple measurements of refractive state can be determined easily and rapidly on a number of different occasions using an autorefractor. In an experimental trial undertaken by Gillan, a 30-year-old female was subjected to 30 autorefractor measurements each taken at various intervals before and after the instillation of Mydriacyl 1% (tropicamide) into her right eye.

The purpose of this paper is to apply multivariate analysis of variance (MANOVA) to Gillan's sample data in order to assess whether instillation of Mydriacyl into the eye affects variability of distance refractive state as well as mean distance refractive state as measured by an autorefractor.

In five of the seven cases where pairwise hypotheses tests were performed, it is concluded that at a 99% level of confidence there is no difference in variability of distance refractive state before and after cycloplegia. In two of the three cases where MANOVA was applied, there is a significant difference at a 95% and at a 99% level of confidence in both variability of distance refractive state and mean distance refractive state with and without cycloplegia.

Keywords: Multivariate analysis of variance (MANOVA), hypothesis testing.

Introduction

The optometrist engaged in research investigates anything that has to do with vision. Different types of refractive variation have been found when measuring refractive state using an autorefractor. Repeated measurements of refractive state reveal variability of the refraction. A cycloplegic refraction is the procedure whereby an individual's refractive error is determined while the muscles that control accommodation are paralysed with cycloplegic agents. Although cycloplegic testing is not usually performed with adult subjects, those who overfocus or underfocus could benefit.

Refractive variability under cycloplegia in a 30-year-old female was considered by Gillan[1]. Analysis of the experimental data was performed by means of multivariate statistical methods developed by Harris[2] and software developed by Harris and Malan. Statistical analysis of refractive variability with small samples was questioned by Malan[3].

This paper applies multivariate analysis of variance (MANOVA) to sample data[1] in order to investigate whether the instillation of a cycloplegic into the right eye of a 30-year-old female subject would affect the variability of her distance refractive state as well as her mean distance refractive state as measured by an autorefractor. This research originates from Gillan's reply[4] to Malan[3]. The method of contrasts as discussed by Abelman[5] and Lemmer[6] is applicable to means only. In this paper a statistical method that none of Gillan[1], Abelman[5] or Lemmer[6] has applied to this sample of optometric data, is considered. Three examples illustrate the method.

*PhD

Received 15 February 2006: revised version accepted 12 June 2006

Examples

The theory associated with the three examples can be found in Harris[2]. The numerical computations are performed using Matlab®.

Let $\sum_1$ be the variance-covariance matrix of the mean distance refractive state of the right eye of a 30-year-old female subject 30 minutes before the instillation of Mydriacyl, $\sum_2$ the variance-covariance matrix of the mean distance refractive state of the eye just prior to instillation, $\sum_3$, $\sum_4$ and $\sum_5$ the variance-covariance matrix of the mean distance refractive state of the eye 15 minutes, 30 minutes and 60 minutes respectively post instillation.

Table 1: Data modified from Gillan[1]. The variance-covariance matrix for vector **h** (vector **h** is indicated in Table 4) is shown for each data set collected. BM1: Data collected 30 minutes prior to instillation of Mydriacyl into the right eye of the subject, BM2: Data collected just prior to instillation, AM1: Data collected 15 minutes post instillation, AM2: Data collected 30 minutes post instillation and AM3: Data collected 60 minutes post instillation. All quantities have units D^2.

BM1	0.00196	0.00045	0.00082
	0.00045	0.00078	0.00003
	0.00082	0.00003	0.00129
BM2	0.00213	–0.00042	0.00129
	–0.00042	0.00094	–0.00017
	0.00129	–0.00017	0.00348
AM1	0.00191	–0.00100	0.00180
	–0.00100	0.00117	–0.00092
	0.00180	–0.00092	0.00328
AM2	0.00170	–0.00118	0.00175
	–0.00118	0.00191	–0.00113
	0.00175	–0.00113	0.00293
AM3	0.00208	–0.00037	0.00201
	–0.00037	0.00103	–0.00074
	0.00201	–0.00074	0.00419

Example 1: Variance-covariances – testing the data from Table 1.

Performing pairwise hypotheses tests is essential for comparison with the MANOVA discussed in example 2. Hypotheses tests are performed at a 5% and at a 1% level of significance. Two possible starting reference values namely $\sum_1$ and $\sum_2$ can be used for performing hypotheses tests. In *AA1* to *AA4* below, $\sum_2$ is used as reference value (same as Gillan[1]), while in *BB1* to *BB4* below, $\sum_1$ is used as reference value. Note that the tests *AA1* and *BB1* are identical.

AA1 $H_0: \sum_2 = \sum_1$ $H_1: \sum_2 \neq \sum_1$

AA2 $H_0: \sum_2 = \sum_3$ $H_1: \sum_2 \neq \sum_3$

AA3 $H_0: \sum_2 = \sum_4$ $H_1: \sum_2 \neq \sum_4$

AA4 $H_0: \sum_2 = \sum_5$ $H_1: \sum_2 \neq \sum_5$

BB1 $H_0: \sum_1 = \sum_2$ $H_1: \sum_1 \neq \sum_2$

BB2 $H_0: \sum_1 = \sum_3$ $H_1: \sum_1 \neq \sum_3$

BB3 $H_0: \sum_1 = \sum_4$ $H_1: \sum_1 \neq \sum_4$

BB4 $H_0: \sum_1 = \sum_5$ $H_1: \sum_1 \neq \sum_5$

The results are presented in Table 2 and are comparable with those of Gillan[1] for the tests *AA1* to *AA4* performed at a 1% level of significance.

Table 2: Test statistics for hypotheses tests on variance-covariance matrices for mean distance refractive state using $\sum_2$ as reference value (same as Gillan[1]), that is, tests *AA1* to *AA4*, as well as test statistics for hypotheses tests on variance-covariance matrices for mean distance refractive state using $\sum_1$ as reference value, that is, tests *BB1* to *BB4*. Note that the tests *AA1* and *BB1* are identical. The value of k is 2. The respective null hypotheses (H_0) are that the variance-covariance matrices for the mean distance refractive states are equal. Critical values are $\chi^2_{0.05,\,6} = 12.592$ and $\chi^2_{0.01,\,6} = 16.812$

Hypothesis Test	Test statistic using equation 48[2], with decision on H_0 denoted by * and **
AA1 and *BB1*	15.3808*
AA2	7.8579
AA3	14.6020*
AA4	3.3291
BB2	24.4964* and **
BB3	28.2304* and **
BB4	13.7980*

* Reject H_0 at 5% level of significance.
** Reject H_0 at 1% level of significance.
No asterisks: Retain H_0 at the appropriate levels of significance.

Example 2: Variance-covariances – testing the data from Table 1 using MANOVA.

Hypotheses tests are performed at a 5% and at a 1% level of significance. The hypotheses to be tested are:

CC $H_0: \sum_1 = \sum_2 = \sum_3 = \sum_4 = \sum_5$
H_1: H_0 is not true

DD $H_0: \sum_2 = \sum_3 = \sum_4 = \sum_5$
H_1: H_0 is not true

EE $H_0: \sum_1 = \sum_3 = \sum_4 = \sum_5$
H_1: H_0 is not true

Table 3: Test statistics for hypotheses tests on variance-covariance matrices for mean distance refractive state using MANOVA. The respective null hypotheses are stated in example 2. Critical values are (for *CC*: $\chi^2_{0.05, 24} = 36.415$ and $\chi^2_{0.01, 24} = 42.980$) and (for *DD* and *EE*: $\chi^2_{0.05, 18} = 28.869$ and $\chi^2_{0.01, 18} = 34.805$).

Hypothesis test and value of k in brackets	Test statistic using equation 32^2, with decision on H_0 denoted by * and **
CC(5)	50.9940 * and **
DD(4)	23.9876
EE(4)	43.6402 * and **

* Reject H_0 at 5% level of significance.
** Reject H_0 at 1% level of significance.
No asterisks: Retain H_0 at the appropriate levels of significance.

The results are presented in Table 3.

Example 3: Mean distance refractive state – testing the data from Table 4 using MANOVA.

Hypotheses are tests performed at a 5% and at a 1% level of significance. Let $\boldsymbol{\mu}_1$ be the mean distance refractive state of the right eye of a 30-year-old female subject 30 minutes before the instillation of Mydriacyl, $\boldsymbol{\mu}_2$ the mean distance refractive state of the right eye just prior to instillation, $\boldsymbol{\mu}_3$, $\boldsymbol{\mu}_4$ and $\boldsymbol{\mu}_5$ the mean distance refractive state of the right eye 15 minutes, 30 minutes and 60 minutes respectively post instillation. The hypotheses to be tested are:

C $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3 = \boldsymbol{\mu}_4 = \boldsymbol{\mu}_5$
H_1: H_0 is not true

D $H_0: \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3 = \boldsymbol{\mu}_4 = \boldsymbol{\mu}_5$
H_1: H_0 is not true

E $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_3 = \boldsymbol{\mu}_4 = \boldsymbol{\mu}_5$
H_1: H_0 is not true

The results are presented in Table 5.

Table 5: Test statistics for hypotheses tests on mean distance refractive state using MANOVA. The respective null hypotheses are stated in example 3. Critical values are (for *C*: $\chi^2_{0.05, 12} = 21.026$ and $\chi^2_{0.01, 12} = 26.217$) and (for *D* and *E*: $\chi^2_{0.05, 9} = 6.919$ and $\chi^2_{0.01, 9} = 21.666$).

Hypothesis test and value of k in brackets	Test statistic using equation 32^2, with decision on H_0 denoted by * and **
C(5)	35.0035 * and **
D(4)	31.0766 * and **
E(4)	13.5038

* Reject H_0 at 5% level of significance.
** Reject H_0 at 1% level of significance.
No asterisks: Retain H_0 at the appropriate levels of significance.

Violation of assumptions

Before any application of statistical procedures, the data should be examined to determine

Table 4: The mean distance refractive state for each set of data as measured by Gillan[1]. Sample size is the same (n = 30) in each case. BM1: Data collected 30 minutes prior to instillation of Mydriacyl into the right eye of the subject, BM2: Data collected immediately prior to instillation, AM1: Data collected 15 minutes after instillation, AM2: Data collected 30 minutes after instillation and AM3: Data collected 60 minutes after instillation. All quantities have units D, except for axis which is measured in degrees.

Mean distance refractive state (conventional form)				*Mean distance refractive state as vector* **h**		
	sph	*cyl*	*axis*	h_1	h_2	h_3
BM1	–0.0377	–0.1407	157	–0.0582	–0.0702	–0.1578
BM2	0.0103	–0.1516	161	–0.0049	–0.0644	–0.1261
AM1	–0.0245	–0.1646	158	–0.0476	–0.0808	–0.1661
AM2	–0.0513	–0.1618	159	–0.0716	–0.0759	–0.1927
AM3	–0.0534	–0.1768	162	–0.0693	–0.0715	–0.2143

whether the assumptions of the test statistics are satisfied. As the raw data were not available to the author, the important issue of departures from normality could not be checked. Lemmer[6] who did have access to the raw data, found that the data were fairly normally distributed and an overall test for normality indicated that the normality assumption was complied with fairly well. According to Sharma[7a] violation of the normality assumption does not have an appreciable effect on the Type I error. The F-test regarding the population means requires the variance-covariance matrices to be equal. According to Lemmer[6], the Box M-test showed this not to be true and casts doubt on the validity of the test. However according to Sharma[7b] the level of significance is not appreciably affected by unequal variance-covariance matrices if the cell sizes are equal.

In the univariate case, if the pooled estimate of the sample variances is to estimate the error variance of the total population, it must be assumed that the k samples are drawn from populations with equal variances, that is,

$$\sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2.$$

Because it is virtually impossible to identify all sources of error, this assumption (called homogeneity of variances or homoscedasticity[8, 9]) is often difficult to justify, and violation can have serious effects on the validity of one's inferences if sample sizes differ markedly from group to group. However, the assumption may be violated without serious risk if the number of observations in each group is the same. Heteroscedasticity[8, 9] is caused by nonnormality of one of the variables, an indirect relationship between variables, or the effects of a data transformation. Heteroscedasticity is not fatal to an analysis, the analysis is weakened, not invalidated. Homoscedasticity is detected with scatterplots and heteroscedasticity is rectified through data transformations similar to those used to achieve normality.

It is common to assume multivariate normality if each variable considered separately follows a normal distribution. MANOVA is robust in the face of most violations of this assumption if sample size exceeds 20.

The primary objective of multivariate analysis of variance is to explore comparisons on the mean vectors. One may wish to investigate hypotheses arising in relation to the basic underlying assumptions of the method. One assumption is that of equal within-group variance-covariance matrices – directly analogous to the homogeneity of variance assumption in univariate analysis[10] as previously described. Tests for this exist, but at least for the univariate case, such tests tend to be more sensitive to departures from normality than the basic test of the analysis of variance. This means that the more sensitive screening test may prevent one from carrying out an analysis which would have been relatively acceptable. When the assumption of equal variance-covariance matrices across groups cannot be maintained, the analysis becomes a generalized Fisher-Behrens problem.[11, 12]

Two approximations based on the likelihood ratio criterion are used to test the hypothesis of equality of variance-covariance matrices. For k multivariate normal populations, the null hypothesis is that $\sum_1 = \sum_2 = ... = \sum_k$ (where $\sum_i$ denotes the variance-covariance matrix of the i-th population group) and the alternative is that $\sum_r \neq \sum_s$ for some r and s. The likelihood ratio test statistic is

$$M = \sum_{t=1}^{k} (n_t - 1) \ln(\det \mathbf{S}) - \sum_{t=1}^{k} (n_t - 1) \ln (\det \mathbf{S}_t).$$

S is the pooled sample variance-covariance matrix:

$$\mathbf{S} = [\sum_{t=1}^{k} (n_t - 1)\, \mathbf{S}_t] / \sum_{t=1}^{k} (n_t - 1).$$

The first approximation leads to MC^{-1} following approximately a χ^2 - distribution with $(k - 1)\, p\, (p + 1)/2$ degrees of freedom. Box's scale factor C^{-1} is defined as

$$C^{-1} = 1 - (2p^2 + 3p - 1)/(6(p + 1)\,(k-1)) \left[\sum_{t=1}^{k} (n_t - 1)^{-1} - \left(\sum (n_t - 1)\right)^{-1}\right].$$

Ideally k and p should not exceed 4 or 5, and each n_t should exceed 20. In the examples discussed, the values of the parameters are k = 2, 4, or 5 respectively, p = 3 (a three-dimensional

sample of 30 measurements), $n_1 = n_2 = n_3 = n_4 = n_5 = 30$, with $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, $\mathbf{S}_4$ and $\mathbf{S}_5$ defined in Table 1. The more complicated second approximation leads to an approximate F- distribution. More detailed theoretical discussions can be found elsewhere.[2, 13]

Discussion

Results emerging from the respective statistical hypotheses tests are interesting and somewhat unexpected. In the context of this paper, the term "significant difference" implies the drug had an effect, while the term "no significant difference" implies the drug had no effect.

For example 1, hypotheses tests *BB2* and *BB3* show a significant difference at both levels of significance. Hypotheses tests *AA1, BB1, AA3* and *BB4* show a significant difference at a 5% level of significance, but no significant difference at a 1% level of significance. Tests *AA2* and *AA4* show no significant difference at both levels of significance. The question needs to be asked why the identical tests *AA1* and *BB1* show a significant difference at a 5% level of significance when no drug had yet been instilled into the right eye of the subject. Concentration, attention, other non-visual sensory inputs and motivation could influence the refractive behaviour of the subject, or it could just be due to normal refractive variation. For hypotheses tests *AA4* and *BB4* (60 minutes post instillation), the drug is already wearing off, so that no significant difference at a 1% level of significance is acceptable. At a 1% level of significance, the results for hypotheses tests *AA1* to *AA4* are comparable with those of Gillan[1].

In example 2, considering tests *CC* and *EE*, at least one of the variance-covariance matrices for the mean distance refractive state of the right eye of this 30-year-old female subject differs significantly at a 5% and at a 1% level of significance. One would have to test pairwise individually to determine which ones are significantly different. Pairwise testing was done in example 1. It was found that at a 1% level of significance for hypotheses tests *AA1–AA4*, none of the variance-covariance matrices for the mean distance refractive state of the right eye of this subject was significantly different, while at a 5% level of significance, hypotheses tests *AA1* and *AA3* showed that $\sum_2 \neq \sum_1$ and $\sum_2 \neq \sum_4$. Hypotheses tests *BB1–BB4* showed that at a 5% level of significance all of the variance-covariance matrices for the mean distance refractive state of the right eye of this 30-year-old female were significantly different, while at a 1% level of significance, hypotheses tests *BB2* and *BB3* showed that $\sum_1 \neq \sum_3$ and $\sum_1 \neq \sum_4$. Considering test *DD*, the variance-covariance matrices for the mean distance refractive state of the right eye of this 30-year-old female are not significantly different at both levels of significance.

In example 3, considering tests *C* and *D*, at least one mean distance refractive state of the right eye of this 30-year-old female subject is significantly different at both levels of significance. One would have to test individually pairwise to determine which ones are significantly different. These hypotheses tests are discussed by Abelman[5]. It was found that at a 5% level of significance, the mean distance refractive state of the right eye of this 30-year-old female 30 minutes before the instillation of Mydriacyl was significantly different from the mean distance refractive state of her right eye just prior to instillation, as well as 30 minutes and 60 minutes respectively post instillation. At a 1% level of significance the mean distance refractive state of her right eye 30 minutes before instillation was significantly different from the mean distance refractive state of her right eye just prior to instillation, as well as 60 minutes post instillation. Further, at both levels of significance, the mean distance refractive state of her right eye just prior to instillation was significantly different from the mean distance refractive state of her right eye 15 minutes, 30 minutes and 60 minutes respectively post instillation. Considering test *E*, at both levels of significance the mean distance refractive state of her right eye measured 30 minutes before instillation was not significantly different from the mean distance refractive state measured 15, 30 and 60 minutes respectively post instillation.

The selection of a cycloplegic agent depends on the desired outcome, the characteristics of the subject receiving the drug and the associated risks. A minimum clinical history of each subject should be undertaken in order to avoid potential adverse drug reactions, both systemic and ocular. For example, one of the side effects of Mydriacyl is dry mouth and this could make

the subject very uncomfortable and influence his/her responses.

Considering all of the above results, one has to ask whether Mydriacyl is in fact an effective cycloplegic for paralysis of the ciliary muscle for the duration of the experiment. Cyclogyl would have been more effective, but it takes longer to dissipate.

Acknowledgement

I thank H Abelman, Optometric Science esearch Group, University of Johannesburg and Honorary Research Associate, Centre or Numerical Analysis and Computational Mathematics, School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, for helpful discussions and suggestions.

The referees are gratefully acknowledged and sincerely thanked for their comments and suggestions which led to the improvement of this paper.

References

1. Gillan WDH. Refractive variability under cycloplegia in a single subject. *S Afr Optom* **59** 19-23.
2. Harris WF. Testing hypotheses on dioptric power. *Optom Vis Sci* 1992 **69** 835-845.
3. Malan DJ. Statistical analysis of refractive variability with small samples. *S Afr Optom* 2000 **59** 94-95.
4. Gillan WDH. Author's reply. *S Afr Optom* 2000 **59** 95.
5. Abelman S. Application of statistical contrasts to mean refractive state. *S Afr Optom* 2006 **65** 27-33.
6. Lemmer HH. The use of contrasts in testing the effect of cycloplegia. *S Afr Optom* 2000 **60** 12-13.
7. Sharma S. *Applied Multivariate Techniques*. New York: Wiley, 1996 a: 375, b: 387.
8. Zar JH. *Biostatistical Analysis*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974 131.
9. Sokal RR, Rohlf FJ. *Introduction to Biostatistics*. San Francisco: WH Freeman and Company, 1973 210.
10. Welch BL. The significance of the difference between means when the population variances are unequal. *Biometrika* 1938 **29** 350-362.
11. Best DJ, Rayner JCW. Welch's approximate solution to the Behrens-Fisher problem. *Technometrics* 1987 **29** 205-210.
12. Scheffé H. Practical solutions of the Behrens-Fisher problem. *J Am Statist Assoc* 1970 **332** 1501-1508.
13. Mardia KV, Kent JT, Bibby JM. *Multivariate Analysis*. London: Academic Press, 1979 120-151.