UNIVERSITI SAINS MALAYSIA

First Semester Examination
2011/2012 Academic Session

January 2012

MSG 366 – Multivariate Analysis
[Analisis Multivariat]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of TWENTY NINE pages of printed material before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi DUA PULUH SEMBILAN muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **all ten** [10] questions.

*[**Arahan:** Jawab **semua sepuluh** [10] soalan.]*

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai].

1. Let

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \text{ with eigenvalues of 1, 4 and -2.}$$

(a) From the given information, obtain $tr(\mathbf{A})$ and $|\mathbf{A}|$.

(b) Find the normalized eigenvectors. Hence, state an orthogonal matrix **C** where $\mathbf{C}'\mathbf{AC} = \mathbf{D}$ and **D** is a diagonal eigenvalues matrix.

(c) Is $\mathbf{A} = \mathbf{CDC}'$? Verify your answer.

[25 marks]

1. Biar

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \textit{ dengan nilai eigen 1, 4 dan -2.}$$

(a) Daripada maklumat yang diberi, dapatkan $tr(\mathbf{A})$ *dan* $|\mathbf{A}|$.

(b) Cari vector eigen ternormal. Dengan itu, nyatakan suatu matrik berortogon **C** *yang mana* $\mathbf{C}'\mathbf{AC} = \mathbf{D}$ *dan* **D** *adalah matrik nilai eigen pepenjuru.*

(c) Adakah $\mathbf{A} = \mathbf{CDC}'$*? Tentusahkan jawapan anda.*

[25 markah]

2. The head measurements on first and second children are observed. Define **HL_1** and **HB_1** as the measurements of head length and head breadth on the first child and **HL_2** and **HB_2** for the second child. Parts of the data and the output are shown in **OUTPUT A**.

(a) Partitioning the results according to the child, state the mean vector and covariance matrix. What can you say about the correlation between the children?

(b) Set up two linear combinations that compare head length and head breadth separately. Write them in a matrix form and find the mean vector and covariance matrix for these linear combinations.

[25 marks]

2. *Ukuran kepala bagi anak pertama dan kedua dicerap. Takrif **HL_1** dan **HB_1** sebagai ukuran panjang kepala dan lebar kepala bagi anak pertama dan **HL_2** dan **HB_2** bagi anak kedua. Sebahagian data dan output ditunjukkan dalam **OUTPUT A**.*

   (a) Petakkan keputusan mengikut anak, nyatakan vector min dan matrik kovarians. Apakah yang boleh dikatakan tentang korelasi di antara anak-anak?

   (b) Setkan dua kombinasi linear yang membandingkan panjang kepala dan lebar kepala secara berasingan. Tuliskan dalam bentuk matrik dan cari vector min dan matrik kovarians bagi kombinasi linear ini.

[25 markah]

3. Let $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$ and $\mathbf{X}_5$ be independent and identically distributed random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Given the two linear combinations of random vectors as

$$\tfrac{1}{5}\mathbf{X}_1 + \tfrac{1}{5}\mathbf{X}_2 + \tfrac{1}{5}\mathbf{X}_3 + \tfrac{1}{5}\mathbf{X}_4 + \tfrac{1}{5}\mathbf{X}_5 \quad \text{and} \quad \mathbf{X}_1 - \mathbf{X}_2 + \mathbf{X}_3 - \mathbf{X}_4 + \mathbf{X}_5.$$

   (a) Find the mean vector and covariance matrices for each of them.

   (b) Obtain the covariance between the two linear combinations of random vectors. What can you say about them?

   (c) If each $\mathbf{X}$ has a trivariate normal distribution, what is the distribution of the two linear combinations? In this case, are they independent?

[25 marks]

3. *Biar $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$ dan $\mathbf{X}_5$ sebagai vector rawak tak bersandar dan secaman dengan vector min $\boldsymbol{\mu}$ dan matrik kovarians $\boldsymbol{\Sigma}$. Diberi dua kombinasi linear bagi vector rawak sebagai*

$$\tfrac{1}{5}\mathbf{X}_1 + \tfrac{1}{5}\mathbf{X}_2 + \tfrac{1}{5}\mathbf{X}_3 + \tfrac{1}{5}\mathbf{X}_4 + \tfrac{1}{5}\mathbf{X}_5 \quad \textit{dan} \quad \mathbf{X}_1 - \mathbf{X}_2 + \mathbf{X}_3 - \mathbf{X}_4 + \mathbf{X}_5.$$

   (a) Cari vector min dan matrik kovarians bagi setiap satu.

   (b) Dapatkan kovarians antara dua kombinasi linear bagi vector rawak. Apakah yang boleh dikata mengenainya?

   (c) Jika setiap $\mathbf{X}$ mempunyai suatu taburan trivariat, apakah taburan bagi dua kombinasi linear? Dalam kes ini, adakah mereka tak bersandar?

[25 markah]

4. Suppose $\mathbf{Y}$ is $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where

$$\boldsymbol{\mu} = \begin{pmatrix} -4 \\ 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 8 & 0 & -1 & 0 \\ 0 & 3 & 0 & 2 \\ -1 & 0 & 5 & 0 \\ 0 & 2 & 0 & 7 \end{pmatrix}.$$

(a) State whether the following random variables are independent:

(i) (Y_1, Y_2) and Y_3,

(ii) (Y_1, Y_2) and Y_4,

(iii) (Y_1, Y_3) and (Y_2, Y_4).

(b) What is the distribution of

(i) Y_3,

(ii) $W_1 = 4Y_1 - 2Y_2 + Y_3 - 3Y_4$?

(c) Find the joint distribution of

(i) Y_1 and Y_3,

(ii) W_1, W_2 and W_3 where W_1 as in b(ii), $W_2 = Y_1 + Y_2 + Y_3 + Y_4$ and $W_3 = -2Y_1 + 3Y_2 + Y_3 - 2Y_4$.

(d) What is the distribution of $\mathbf{U} = (\boldsymbol{\Sigma}^{1/2})^{-1}(\mathbf{Y} - \boldsymbol{\mu})$?

[40 marks]

4. Katakan $\mathbf{Y}$ adalah $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ yang mana

$$\boldsymbol{\mu} = \begin{pmatrix} -4 \\ 2 \\ 5 \\ -1 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 8 & 0 & -1 & 0 \\ 0 & 3 & 0 & 2 \\ -1 & 0 & 5 & 0 \\ 0 & 2 & 0 & 7 \end{pmatrix}.$$

(a) Nyatakan sama ada pemboleh ubah rawak berikut adalah tak bersandar:

(i) (Y_1, Y_2) *dan* Y_3,

(ii) (Y_1, Y_2) *dan* Y_4,

(iii) (Y_1, Y_3) *dan* (Y_2, Y_4).

(b) Apakah taburan bagi

(i) Y_3,

(ii) $W_1 = 4Y_1 - 2Y_2 + Y_3 - 3Y_4$ *?*

(d) *Cari taburan tercantum bagi*

(i) Y_1 *dan* Y_3,

(ii) W_1, W_2 *dan* W_3 *yang mana* W_1 *adalah seperti di b(ii),* $W_2 = Y_1 + Y_2 + Y_3 + Y_4$ *dan* $W_3 = -2Y_1 + 3Y_2 + Y_3 - 2Y_4$.

(d) *Apakah taburan bagi* $\mathbf{U} = (\mathbf{\Sigma}^{1/2})^{-1}(\mathbf{Y} - \boldsymbol{\mu})$ *?*

[40 markah]

5. Write a paragraph on detecting outliers.

[10 marks]

5. Tuliskan suatu perenggan tentang cara mendapati cerapan terpencil.

[10 markah]

6. The data and output in **OUTPUT B** consist of measurements y_1, y_2, y_3 and y_4 of the ramus bone length (RBL) at four different ages (8 years, 8.5 years, 9 years and 9.5 years) on each of 20 boys.

(a) What is $\bar{\mathbf{y}}, \mathbf{S}$ and $\mathbf{R}$?

(b) Test $H_0 : \boldsymbol{\mu} = (48, 49, 50, 51)'$. Use $\alpha = 0.05$.

(c) Obtain a 95% Bonferroni simultaneous confidence interval for the means. Based on these intervals, what can you say about the test in (b)?

(d) Explain how you can test $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$. Do not perform the test.

[40 marks]

6. Data dan output dalam ***OUTPUT B*** *terdiri dari ukuran* y_1, y_2, y_3 *dan* y_4 *bagi panjang tulang 'ramus' (RBL) pada empat umur berbeza (8 tahun, 8.5 tahun, 9 tahun and 9.5 tahun) ke atas setiap 20 kanak-kanak lelaki.*

(a) *Apakah* $\bar{\mathbf{y}}, \mathbf{S}$ *dan* $\mathbf{R}$ *?*

(b) *Uji* $H_0 : \boldsymbol{\mu} = (48, 49, 50, 51)'$. *Guna* $\alpha = 0.05$.

(c) *Dapatkan suatu selang keyakinan serentak 95% Bonferroni bagi min. Berdasarkan selang ini, apakah yang boleh anda katakan tentang ujian dalam (b)?*

(d) *Terangkan bagimana anda boleh uji* $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$. *Jangan jalankan ujian ini.*

[40 markah]

7. The data in the table below are speed of calculation from a design with two factors. Factor A is a comparison of two tasks and factor B is a comparison of two types of calculators.

Subjects	A_1		A_2	
	B_1	B_2	B_1	B_2
S1	30	21	21	14
S2	22	13	22	5
S3	29	13	18	17
S4	12	7	16	14
S5	23	24	23	8

(a) Identify the design.

(b) Given

$$\bar{\mathbf{y}} = \begin{pmatrix} 23.2 \\ 15.6 \\ 20.0 \\ 11.6 \end{pmatrix}, \quad \mathbf{S} = \begin{bmatrix} 51.7 & & & \\ 29.8 & 46.8 & & \\ 9.2 & 16.2 & 8.5 & \\ 7.4 & -8.7 & -10.5 & 24.3 \end{bmatrix} \quad \text{and}$$

$$(\mathbf{CSC})^{-1} = 10^{-2} \begin{bmatrix} 1.2 & 0.9 & 0.3 \\ 0.9 & 2.9 & 0 \\ 0.3 & 0 & 0.8 \end{bmatrix}.$$

State the contrast to test $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$. Perform the test at $\alpha = 0.05$. Give your conclusion.

(c) Construct the 95% simultaneous confidence intervals for the contrast used in (b). Give your conclusion.

[40 marks]

7. *Data dalam jadual di bawah adalah kelajuan kalkulator dari suatu rekabentuk dengan dua faktor. Faktor A adalah perbandingan dua tugas dan faktor B adalah perbandingan dua jenis kalkulator.*

Subjek	A_1		A_2	
	B_1	B_2	B_1	B_2
S1	*30*	*21*	*21*	*14*
S2	*22*	*13*	*22*	*5*
S3	*29*	*13*	*18*	*17*
S4	*12*	*7*	*16*	*14*

S5	*23*	*24*	*23*	*8*

(a) *Camkan rekabentuk ini.*

(b) *Diberi*

$$\bar{\mathbf{y}} = \begin{pmatrix} 23.2 \\ 15.6 \\ 20.0 \\ 11.6 \end{pmatrix}, \quad \mathbf{S} = \begin{bmatrix} 51.7 & & & \\ 29.8 & 46.8 & & \\ 9.2 & 16.2 & 8.5 & \\ 7.4 & -8.7 & -10.5 & 24.3 \end{bmatrix} \quad \textit{dan}$$

$$(\mathbf{CSC})^{-1} = 10^{-2}\begin{bmatrix} 1.2 & 0.9 & 0.3 \\ 0.9 & 2.9 & 0 \\ 0.3 & 0 & 0.8 \end{bmatrix}.$$

Nyatakan kontras untuk menguji $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$. *Jalankan ujian pada* $\alpha = 0.05$. *Beri kesimpulan anda.*

(c) *Bina selang keyakinan serentak 95% bagi kontras yang digunakan dalam (b). Beri kesimpulan anda.*

[40 markah]

8. Judges scores on fish prepared by three methods are noted. Twelve fish were cooked by each method, and several judges tasted fish samples of the fish dishes and rated each on four variables: y_1 = aroma, y_2 = flavour, y_3 = texture, and y_4 = moisture. The partial data and the output are provided by **OUTPUT C**. Interpret the results and give your conclusions.

[30 marks]

8. *Markah juri ke atas ikan yang dimasak dengan tiga cara dicatat. Dua belas ikan dimasak dengan setiap cara, dan beberapa juri merasai sampel ikan yang dimasak dan menilai setiap satu berdasarkan empat pemboleh ubah:* y_1 *= aroma,* y_2 *= rasa,* y_3 *='texture', dan* y_4 *= kelembapan. Sebahagian data dan output diberi oleh* ***OUTPUT C****. Tafsir keputusan dan beri kesimpulan anda.*

[30 markah]

9. A discriminant analysis is also performed on the data in question 8. The output from the analysis and MANOVA are as in **OUTPUT D**.

(a) Interpret the results and give your conclusion.

(b) Can we perform a cluster analysis on the data? Explain.

[30 marks]

*9. Suatu analisis pembezalayan juga dijalankan ke atas data dalam soalan 8. Output dari analisis dan MANOVA adalah seperti dalam **OUTPUT D**.*

(a) Tafsir keputusan dan beri kesimpulan anda.

(b) Bolehkah kita jalankan suatu analisis kelompok ke atas data ini? Terangkan.

[30 markah]

10. Combining the three groups into a single sample, the data in question 8 are analyzed. The SPSS factor analysis program was used to fit the factor models. The output is as in **OUTPUT E**. Interpret the results.

[35 marks]

*10. Dengan menggabungkan tiga kumpulan ke dalam suatu sampel tunggal, data dalam soalan 8 dianalisa. Program analisis faktor SPSS diguna untuk penyuaian model faktor. Output adalah seperti dalam **OUTPUT E**. Tafsir keputusan.*

[35 markah]

APPENDIX / LAMPIRAN

FORMULAE

The notations are as given in the lectures.

1. Suppose $\mathbf{X}$ has $E(\mathbf{X}) = \boldsymbol{\mu}$ and $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$. Thus $\mathbf{c}'\mathbf{X}$ has mean, $\mathbf{c}'\boldsymbol{\mu}$, and variance, $\mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$.

2. Bivariate normal p.d.f:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \times \exp\left\{-\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}$$

3. Multivariate normal p.d.f:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \mathbf{e}^{-1/2(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

4. If $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then
   (a) $\mathbf{aX} \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$
   (b) $\mathbf{AX} \sim N_q(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$
   (c) $\mathbf{X}+\mathbf{d} \sim N_p(\boldsymbol{\mu}+\mathbf{d}, \boldsymbol{\Sigma})$, $\mathbf{d}$ is a vector of constant
   (d) $(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \sim \chi_p^2$

5. Let $\mathbf{X}_j \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$, $j = 1, \ldots, n$ be mutually independent with the same covariance matrix $\boldsymbol{\Sigma}$. Then $\mathbf{V}_1 = \sum_{j=1}^{n} c_j \mathbf{X}_j \sim N_p\left(\sum_{j=1}^{n} c_i \boldsymbol{\mu}_j, \left(\sum_{j=1}^{n} c_j^2\right)\boldsymbol{\Sigma}\right)$. Moreover, $\mathbf{V}_1$ and $\mathbf{V}_2 = \sum_{j=1}^{n} b_j \mathbf{X}_j$ are jointly multivariate normal with covariance matrix

$$\begin{bmatrix} \left(\sum_{j=1}^{n} c_j^2\right)\boldsymbol{\Sigma} & (\mathbf{b}'\mathbf{c})\boldsymbol{\Sigma} \\ (\mathbf{b}'\mathbf{c})\boldsymbol{\Sigma} & \left(\sum_{j=1}^{n} b_j^2\right)\boldsymbol{\Sigma} \end{bmatrix}.$$

6. If $\mathbf{A}_1 \sim W_{m_1}(\mathbf{A}_1 | \boldsymbol{\Sigma})$ independently of $\mathbf{A}_2$, which $\mathbf{A}_2 \sim W_{m_2}(\mathbf{A}_2 | \boldsymbol{\Sigma})$, then $\mathbf{A}_1 + \mathbf{A}_2 \sim W_{m_1+m_2}(\mathbf{A}_1 + \mathbf{A}_2 | \boldsymbol{\Sigma})$. Also, if $\mathbf{A} \sim W_m(\mathbf{A} | \boldsymbol{\Sigma})$, then

$$\mathbf{CAC}' \sim W_m(\mathbf{CAC}' | \mathbf{C\Sigma C}').$$

7. One-sample :

(a) $$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu})$$

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{X}_j, \quad \mathbf{S} = \frac{1}{n-1}\sum_{j-1}^{n} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$$

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p}$$

(b) $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{a}'\boldsymbol{\mu}$:

$$\mathbf{a}'\bar{\mathbf{X}} \pm \sqrt{\frac{p(n-1)}{n-p} F_{p.n-p}(\alpha)\left(\frac{\mathbf{a}'\mathbf{S}\,\mathbf{a}}{n}\right)}$$

(c) Bonferroni $100(1-\alpha)\%$ confidence interval for $\mathbf{a}'\boldsymbol{\mu}$:

$$\mathbf{a}'\bar{\mathbf{X}} \pm t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{\mathbf{a}'\mathbf{S}\,\mathbf{a}}{n}}$$

(d) $100(1-\alpha)\%$ large sample confidence interval for $\mathbf{a}'\boldsymbol{\mu}$:

$$\mathbf{a}'\bar{\mathbf{X}} \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{\mathbf{a}'\mathbf{S}\mathbf{a}}{n}}$$

(e) Wilk Lambda, $\Lambda^{2/n} = \left(\frac{\left|\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\mathbf{x}_j-\bar{\mathbf{x}})'\right|}{\left|\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)(\mathbf{x}_j-\boldsymbol{\mu}_0)'\right|} \right) = \left(1+\frac{T^2}{(n-1)}\right)^{-1}$

8. Paired comparisons

(a) $T^2 = n(\bar{\mathbf{D}}-\boldsymbol{\delta})'\boldsymbol{\delta}_d^{-1}(\bar{\mathbf{D}}-\boldsymbol{\delta})$

$$\bar{\mathbf{D}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{D}_j \qquad \mathbf{S}_d = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{D}_j-\bar{\mathbf{D}})(\mathbf{D}_j-\bar{\mathbf{D}})'$$

$$T^2 \sim \left[\frac{(n-1)p}{(n-p)}\right]F_{p,n-p}$$

(b) $100(1-\alpha)\%$ simultaneous confidence interval for δ_i:

$$\bar{d}_i \pm \sqrt{\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{d_i}^2}{n}}$$

$\bar{d}_i = i^{th}$ element of $\bar{\mathbf{d}}$

$s_{d_i}^2 = i^{th}$ diagonal element of $\mathbf{S}_d$

(c) Bonferroni $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{c}'\boldsymbol{\mu}$:

$$\bar{d}_i \pm t_{(n-1)}(\alpha/2p)\sqrt{\frac{s_{d_i}^2}{n}}$$

9. Repeated Measure Design

(a) Let $\mathbf{C}$ be a contrast matrix

$$T^2 = n\,(\mathbf{C}\bar{\mathbf{x}})'\,(\mathbf{CSC}')^{-1}\,\mathbf{C}\bar{\mathbf{x}}$$

$$T^2 \sim \frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,\,n-q+1}(\alpha)$$

(b) $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{c}'\boldsymbol{\mu}$:

$$\mathbf{c}'\bar{\mathbf{x}} \pm \sqrt{\frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(\alpha)}\sqrt{\frac{\mathbf{c}'\mathbf{Sc}}{n}}$$

10. Two independent samples:

(a) $$T^2 = \left[\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right]'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_p\right]^{-1}\left[\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right]$$

$$T^2 \sim \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p,n_1+n_2-p-1}$$

$$\mathbf{S}_p = \frac{n_i-1}{n_i+n_2-2}\mathbf{S}_1 + \frac{n_2-1}{n_1+n_2-2}\mathbf{S}_2$$

$$\mathbf{S}_i = \frac{\sum_{j=1}^{n_i}(\mathbf{x}_{ij}-\bar{\mathbf{x}}_i)(\mathbf{x}_{ij}-\bar{\mathbf{x}}_i)'}{n_i-1}$$

(b) $100(1-\alpha)\%$ simultaneous confidence intervals for $\mathbf{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$:

$$\mathbf{a}'(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) \pm c\sqrt{\mathbf{a}'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_p\,\mathbf{a}}$$

where $c^2 = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p,n_1+n_2-p-1}(\alpha)$

(c) For large $n_1 - p,$ and $n_2 - p,$ $100(1-\alpha)\%$ simultaneous confidence interval for $\mathbf{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$:

$$\mathbf{a}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \pm c\sqrt{\mathbf{a}'\left(\frac{1}{n_1}\mathbf{S}_1+\frac{1}{n_2}\mathbf{S}_2\right)\mathbf{a}}$$

where $c^2 = \chi^2_p(\alpha)$

11. One-way MANOVA:

(a)
$$\mathbf{B} = \sum_{\ell=1}^{g} n_\ell (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$$
$$\mathbf{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$$
$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|}$$

(b) Bartlett: If $\sum n_\ell = n$ is large,

$$-\left(n-1-\frac{p+g}{2}\right)\ln \Lambda^* = -\left((n-1) - \frac{p+g}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|}\right) \sim \chi^2_{p(g-1)}$$

(c) $100(1-\alpha)\%$ simultaneous confidence intervals for $\tau_{ki} - \tau_{\ell i}$:

$$\bar{x}_{ki} - \bar{x}_{\ell i} \pm t_{n-g}\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{w_{ii}}{n-g}\left(\frac{1}{n_k}+\frac{1}{n_\ell}\right)}, \quad n = \sum_{\ell=1}^{g} n_\ell$$
$$i = 1, 2, \ldots, p, \quad \ell < k = 1, 2, \ldots, g$$
$w_{ii} = i^{th}$ diagonal element of $\mathbf{W}$.

OUTPUT A

Row	HL_1	HB_1	HL_2	HB_2
1	191	155	179	145
2	195	149	201	152
3	181	148	185	149
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
21	192	154	185	152
22	174	143	178	147
23	176	139	176	143
24	197	167	200	158
25	190	163	187	150

Descriptive Statistics: HL_1, HB_1, HL_2, HB_2

Variable	N	Mean	Median	TrMean	StDev	SE Mean
HL_1	25	185.72	188.00	185.74	9.76	1.95
HB_1	25	151.12	151.00	151.04	7.37	1.47
HL_2	25	183.84	185.00	184.09	10.04	2.01
HB_2	25	149.24	149.00	149.65	6.71	1.34

Covariances: HL_1, HB_1, HL_2, HB_2

	HL_1	HB_1	HL_2	HB_2
HL_1	95.2933			
HB_1	52.8683	54.3600		
HL_2	69.6617	51.3117	100.8067	
HB_2	46.1117	35.0533	56.5400	45.0233

OUTPUT B

Row	y1	y2	y3	y4
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.4
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	48.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8

Descriptive Statistics: y1, y2, y3, y4

Variable	N	Mean	Median	TrMean	StDev	SE Mean
y1	20	48.7	48.150	48.600	2.516	0.563
y2	20	49.6	49.050	49.589	2.540	0.568
y3	20	50.6	50.750	50.567	2.630	0.588
y4	20	51.4	51.800	51.456	2.726	0.610

Matrix COVA1

6.3	6.2	5.8	5.5
6.2	6.4	6.2	5.9
5.8	6.2	6.9	6.9
5.5	5.9	6.9	7.4

Matrix InCOVA1

2.7	-2.9	0.5	-0.2
-2.9	4.4	-2.2	0.8
0.5	-2.2	4.6	-2.9
-0.2	0.8	-2.9	2.3

OUTPUT C

```
Row    Method     y1     y2     y3     y4

  1      1       5.4    6.0    6.3    6.7
  2      1       5.2    6.5    6.0    5.8
  3      1       6.1    5.9    6.0    7.0
  4      1       4.8    5.0    4.9    5.0
  .      .        .      .      .      .
  .      .        .      .      .      .
  .      .        .      .      .      .

 32      3       4.8    4.6    5.7    5.7
 33      3       5.3    5.4    6.8    6.6
 34      3       4.6    4.4    5.7    5.6
 35      3       4.5    4.0    5.0    5.9
 36      3       4.4    4.2    5.6    5.5
```

Descriptive Statistics

	Method	Mean	Std. Deviation	N
y1	1	5.383	.5844	12
	2	5.258	.7633	12
	3	4.975	.5429	12
	Total	5.206	.6427	36
y2	1	5.733	.4793	12
	2	5.233	.5742	12
	3	4.833	.4599	12
	Total	5.267	.6178	36
y3	1	5.442	.6640	12
	2	5.308	.5946	12
	3	5.908	.5107	12
	Total	5.553	.6322	36
y4	1	5.983	.6965	12
	2	5.875	.5172	12
	3	6.233	.4559	12
	Total	6.031	.5701	36

Box's Test of Equality of Covariance Matrices[a]

Box's M	17.212
F	.707
df1	20
df2	3909.028
Sig.	.823

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.993	1103.356[a]	4.000	30.000	.000
	Wilks' Lambda	.007	1103.356[a]	4.000	30.000	.000
	Hotelling's Trace	147.114	1103.356[a]	4.000	30.000	.000
	Roy's Largest Root	147.114	1103.356[a]	4.000	30.000	.000
Method	Pillai's Trace	.860	5.845	8.000	62.000	.000
	Wilks' Lambda	.224	8.329[a]	8.000	60.000	.000
	Hotelling's Trace	3.079	11.161	8.000	58.000	.000
	Roy's Largest Root	2.951	22.874[b]	4.000	31.000	.000

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c. Design: Intercept + Method

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
y1	.684	2	33	.512
y2	.080	2	33	.923
y3	.592	2	33	.559
y4	1.167	2	33	.324

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Method

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	y1	1.051[a]	2	.525	1.293	.288
	y2	4.880[b]	2	2.440	9.495	.001
	y3	2.382[c]	2	1.191	3.386	.046
	y4	.811[d]	2	.405	1.266	.295
Intercept	y1	975.521	1	975.521	2400.910	.000
	y2	998.560	1	998.560	3885.906	.000
	y3	1110.000	1	1110.000	3155.719	.000
	y4	1309.234	1	1309.234	4089.096	.000

Method	y1	1.051	2	.525	1.293	.288
	y2	4.880	2	2.440	9.495	.001
	y3	2.382	2	1.191	3.386	.046
	y4	.811	2	.405	1.266	.295
Error	y1	13.408	33	.406		
	y2	8.480	33	.257		
	y3	11.607	33	.352		
	y4	10.566	33	.320		
Total	y1	989.980	36			
	y2	1011.920	36			
	y3	1123.990	36			
	y4	1320.610	36			
Corrected Total	y1	14.459	35			
	y2	13.360	35			
	y3	13.990	35			
	y4	11.376	35			

a. R Squared = .073 (Adjusted R Squared = .016)

b. R Squared = .365 (Adjusted R Squared = .327)

c. R Squared = .170 (Adjusted R Squared = .120)

d. R Squared = .071 (Adjusted R Squared = .015)

Between-Subjects SSCP Matrix

			y1	y2	y3	y4
Hypothesis	Intercept	y1	975.521	986.973	1040.591	1130.126
		y2	986.973	998.560	1052.807	1143.393
		y3	1040.591	1052.807	1110.000	1205.508
		y4	1130.126	1143.393	1205.508	1309.234
	Method	y1	1.051	2.173	-1.376	-.760
		y2	2.173	4.880	-2.373	-1.257
		y3	-1.376	-2.373	2.382	1.384
		y4	-.760	-1.257	1.384	.811
Error		y1	13.408	7.723	8.675	5.864
		y2	7.723	8.480	7.527	6.213
		y3	8.675	7.527	11.607	7.038
		y4	5.864	6.213	7.038	10.566

Based on Type III Sum of Squares

Multiple Comparisons

Bonferroni

Dependent Variable	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
y1	1	2	.125	.2602	1.000	-.531	.781
		3	.408	.2602	.378	-.248	1.065
	2	1	-.125	.2602	1.000	-.781	.531
		3	.283	.2602	.852	-.373	.940
	3	1	-.408	.2602	.378	-1.065	.248
		2	-.283	.2602	.852	-.940	.373
y2	1	2	Double-click to activate	.2069	.064	-.022	1.022
		3		.2069	.000	.378	1.422
	2	1	-.500	.2069	.064	-1.022	.022
		3	.400	.2069	.186	-.122	.922
	3	1	-.900*	.2069	.000	-1.422	-.378
		2	-.400	.2069	.186	-.922	.122
y3	1	2	.133	.2421	1.000	-.477	.744
		3	-.467	.2421	.188	-1.077	.144
	2	1	-.133	.2421	1.000	-.744	.477
		3	-.600	.2421	.055	-1.211	.011
	3	1	.467	.2421	.188	-.144	1.077
		2	.600	.2421	.055	-.011	1.211
y4	1	2	.108	.2310	1.000	-.474	.691
		3	-.250	.2310	.861	-.833	.333
	2	1	-.108	.2310	1.000	-.691	.474
		3	-.358	.2310	.391	-.941	.224

OUTPUT D

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
y1	.927	1.293	2	33	.288
y2	.635	9.495	2	33	.001
y3	.830	3.386	2	33	.046
y4	.929	1.266	2	33	.295

Covariance Matrices[a]

Method		y1	y2	y3	y4
1	y1	.342	.149	.262	.231
	y2	.149	.230	.245	.223
	y3	.262	.245	.441	.327
	y4	.231	.223	.327	.485
2	y1	.583	.359	.372	.192
	y2	.359	.330	.250	.178
	y3	.372	.250	.354	.161
	y4	.192	.178	.161	.267
3	y1	.295	.195	.155	.111
	y2	.195	.212	.190	.163
	y3	.155	.190	.261	.152
	y4	.111	.163	.152	.208
Total	y1	.413	.283	.209	.146
	y2	.283	.382	.147	.142
	y3	.209	.147	.400	.241
	y4	.146	.142	.241	.325

a. The total covariance matrix has 35 degrees of freedom.

Box's Test of Equality of Covariance Matrices

Test Results

Box's M		17.212
F	Approx.	.707
	df1	20
	df2	3909.028
	Sig.	.823

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.951[a]	95.9	95.9	.864
2	.127[a]	4.1	100.0	.336

a. First 2 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.224	47.059	8	.000
2	.887	3.775	3	.287

Canonical Discriminant Function Coefficients

	Function	
	1	2
y1	.119	-1.823
y2	3.064	1.714
y3	-1.992	1.397
y4	-.776	-.151
(Constant)	-1.015	-6.384

Unstandardized coefficients

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
y1	.076	-1.162
y2	1.553	.869
y3	-1.182	.828
y4	-.439	-.085

...22/-

Structure Matrix

	Function	
	1	2
y1	.163*	.001
y3	-.229	.625*
y2	.424	.600*
y4	-.133	.439*

Functions at Group Centroids

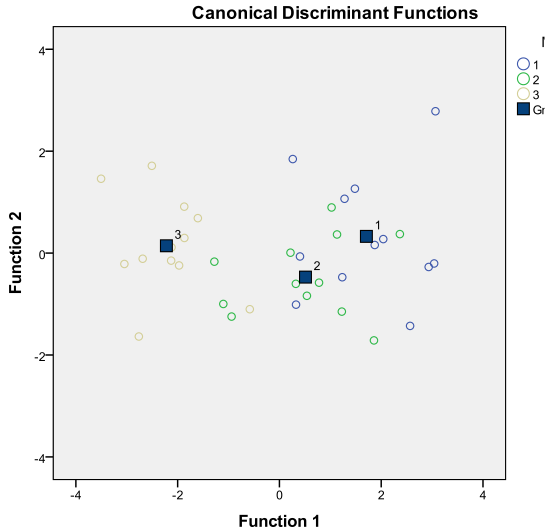
	Function	
Method	1	2
1	1.709	.328
2	.512	-.471
3	-2.221	.144

Unstandardized canonical discriminant functions evaluated at group means

Classification Function Coefficients

	Method		
	1	2	3
y1	.808	2.122	.676
y2	15.151	10.113	2.792
y3	-1.030	.239	6.543
y4	10.015	11.065	13.093
(Constant)	-73.867	-66.279	-69.665

Fisher's linear discriminant functions



Classification Results[b,c]

		Method	Predicted Group Membership 1	2	3	Total
Original	Count	1	9	3	0	12
		2	3	7	2	12
		3	0	1	11	12
	%	1	75.0	25.0	.0	100.0
		2	25.0	58.3	16.7	100.0
		3	.0	8.3	91.7	100.0
Cross-validated[a]	Count	1	7	5	0	12
		2	4	5	3	12
		3	0	1	11	12
	%	1	58.3	41.7	.0	100.0
		2	33.3	41.7	25.0	100.0
		3	.0	8.3	91.7	100.0

OUTPUT E

Correlation Matrix

		y1	y2	y3	y4
Correlation	y1	1.000	.712	.513	.398
	y2	.712	1.000	.377	.402
	y3	.513	.377	1.000	.668
	y4	.398	.402	.668	1.000

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.635
Bartlett's Test of Sphericity	Approx. Chi-Square	54.432
	df	6
	Sig.	.000

Anti-image Matrices

		y1	y2	y3	y4
Anti-image Covariance	y1	.421	-.290	-.155	.035
	y2	-.290	.470	.053	-.106
	y3	-.155	.053	.475	-.297
	y4	.035	-.106	-.297	.525
Anti-image Correlation	y1	.628[a]	-.652	-.346	.075
	y2	-.652	.627[a]	.112	-.214
	y3	-.346	.112	.636[a]	-.595
	y4	.075	-.214	-.595	.654[a]

a. Measures of Sampling Adequacy(MSA)

Communalities

	Initial	Extraction
y1	1.000	.914
y2	1.000	.935
y3	1.000	.950
y4	1.000	.970

...25/-

Communalities

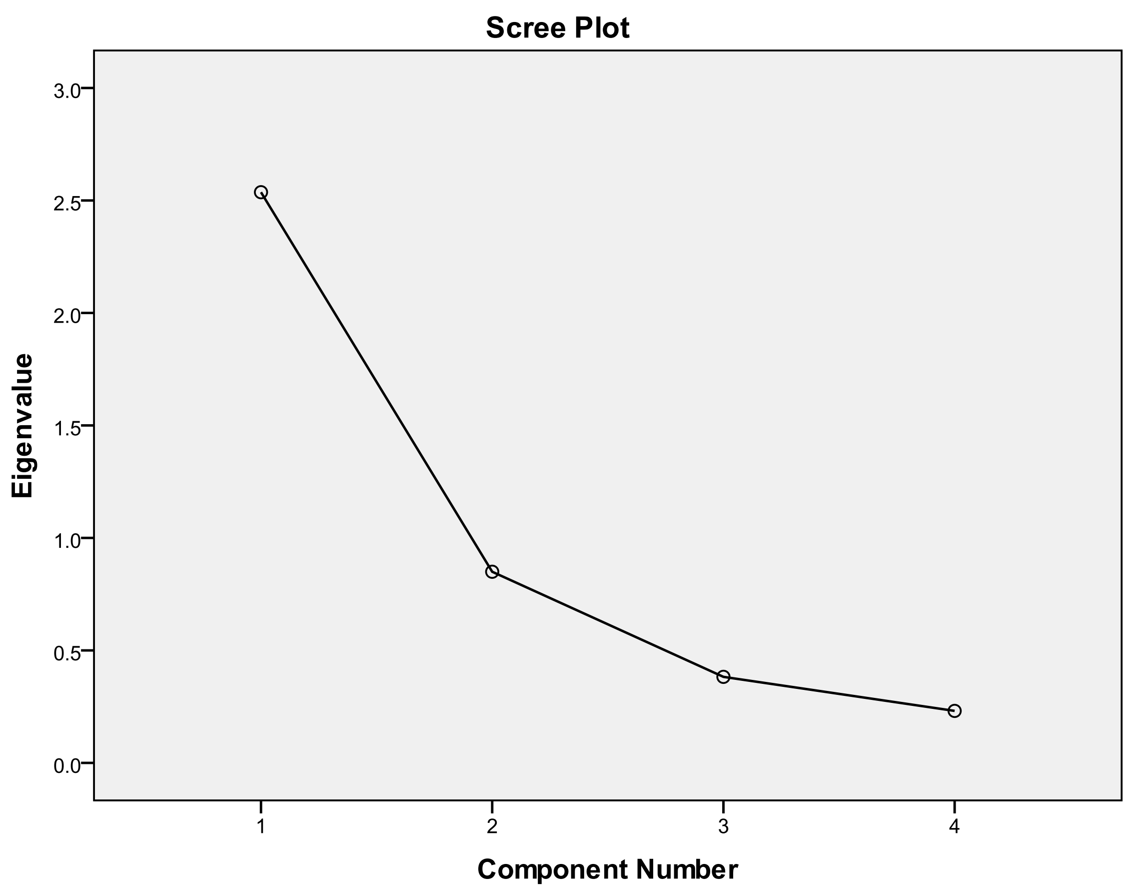
	Initial	Extraction
y1	1.000	.914
y2	1.000	.935
y3	1.000	.950
y4	1.000	.970

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.537	63.417	63.417	2.537	63.417	63.417	1.619	40.475	40.475
2	.850	21.243	84.660	.850	21.243	84.660	1.091	27.285	67.759
3	.382	9.562	94.222	.382	9.562	94.222	1.058	26.462	94.222
4	.231	5.778	100.000						

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component		
	1	2	3
y1	.830	-.403	-.250
y3	.803	.432	-.345
y2	.783	-.504	.263
y4	.769	.497	.363

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Rotated Component Matrix[a]

	Component		
	1	2	3
y2	.925	.282	
y1	.829		.475
y4	.197	.908	.325
y3	.193	.431	.853

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

General Linear Model

Between-Subjects Factors

		N
Method	1	12
	2	12
	3	12

Descriptive Statistics

	Method	Mean	Std. Deviation	N
REGR factor score 1 for analysis 2	1	.7303861	.59468167	12
	2	.1320553	.95821056	12
	3	-.8624414	.70540480	12
	Total	.0000000	1.00000000	36
REGR factor score 2 for analysis 2	1	-.0132627	1.18907407	12
	2	-.3060438	.92361556	12
	3	.3193065	.83741295	12
	Total	.0000000	1.00000000	36
REGR factor score 3 for analysis 2	1	-.4622201	.92650890	12
	2	-.2421315	1.01890497	12
	3	.7043515	.66857513	12
	Total	.0000000	1.00000000	36

Box's Test of Equality of Covariance Matrices[a]

Box's M	8.078
F	.583
df1	12
df2	5277.462
Sig.	.858

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.000	.000[a]	3.000	31.000	1.000
	Wilks' Lambda	1.000	.000[a]	3.000	31.000	1.000
	Hotelling's Trace	.000	.000[a]	3.000	31.000	1.000
	Roy's Largest Root	.000	.000[a]	3.000	31.000	1.000
Method	Pillai's Trace	.774	6.741	6.000	64.000	.000
	Wilks' Lambda	.254	10.156[a]	6.000	62.000	.000
	Hotelling's Trace	2.818	14.092	6.000	60.000	.000
	Roy's Largest Root	2.778	29.629[b]	3.000	32.000	.000

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
REGR factor score 1 for analysis 2	1.679	2	33	.202
REGR factor score 2 for analysis 2	1.085	2	33	.350
REGR factor score 3 for analysis 2	2.082	2	33	.141

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Method

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	REGR factor score 1 for analysis 2	15.536[a]	2	7.768	13.171	.000
	REGR factor score 2 for analysis 2	2.350[b]	2	1.175	1.187	.318
	REGR factor score 3 for analysis 2	9.221[c]	2	4.610	5.902	.006
Intercept	REGR factor score 1 for analysis 2	.000	1	.000	.000	1.000
	REGR factor score 2 for analysis 2	.000	1	.000	.000	1.000
	REGR factor score 3 for analysis 2	.000	1	.000	.000	1.000
Method	REGR factor score 1 for analysis 2	15.536	2	7.768	13.171	.000
	REGR factor score 2 for analysis 2	2.350	2	1.175	1.187	.318
	REGR factor score 3 for analysis 2	9.221	2	4.610	5.902	.006
Error	REGR factor score 1 for analysis 2	19.464	33	.590		
	REGR factor score 2 for analysis 2	32.650	33	.989		
	REGR factor score 3 for analysis 2	25.779	33	.781		

Total	REGR factor score 1 for analysis 2	35.000	36			
	REGR factor score 2 for analysis 2	35.000	36			
	REGR factor score 3 for analysis 2	35.000	36			
Corrected Total	REGR factor score 1 for analysis 2	35.000	35			
	REGR factor score 2 for analysis 2	35.000	35			
	REGR factor score 3 for analysis 2	35.000	35			

a. R Squared = .444 (Adjusted R Squared = .410)

b. R Squared = .067 (Adjusted R Squared = .011)

c. R Squared = .263 (Adjusted R Squared = .219)

Pairwise Comparisons

Dependent Variable	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
REGR factor score 1 for analysis 2	1	2	.598	.314	.195	-.192	1.389
		3	1.593*	.314	.000	.802	2.384
	2	1	-.598	.314	.195	-1.389	.192
		3	.994*	.314	.010	.204	1.785
	3	1	-1.593*	.314	.000	-2.384	-.802
		2	-.994*	.314	.010	-1.785	-.204
REGR factor score 2 for analysis 2	1	2	.293	.406	1.000	-.731	1.317
		3	-.333	.406	1.000	-1.357	.692
	2	1	Double-click to activate	.406	1.000	-1.317	.731
		3	-.625	.406	.399	-1.650	.399
	3	1	.333	.406	1.000	-.692	1.357
		2	.625	.406	.399	-.399	1.650
REGR factor score 3 for analysis 2	1	2	-.220	.361	1.000	-1.130	.690
		3	-1.167*	.361	.008	-2.077	-.256
	2	1	.220	.361	1.000	-.690	1.130
		3	-.946*	.361	.039	-1.857	-.036
	3	1	1.167*	.361	.008	.256	2.077
		2	.946*	.361	.039	.036	1.857

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

- ooo O ooo -