MULTIVARIATE LINEAR CALIBRATION FOR LINEAR COMBINATIONS

Nakao, Hiroyuki
Department of Public Health, Miyazaki Medical College

https://doi.org/10.5109/13508

出版情報：Bulletin of informatics and cybernetics. 34 (1), pp.51-60, 2002-10. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 34, No. 1, 2002

MULTIVARIATE LINEAR CALIBRATION FOR LINEAR COMBINATIONS

By

Hiroyuki NAKAO*

Abstract

This paper discusses confidence regions of linear combinations of explanatory variables in multivariate linear calibration. Approximated confidence regions are considered, since the exact confidence regions in multivariate linear calibration are often empty or are not bounded. A numerical example is given for the approximated confidence regions.

Key Words and Phrases: Calibration, Confidence Region, Multivariate Normal Distribution, Linear Combination.

1. Introduction

In a multivariate regression model, we are interested in constructing confidence regions of the explanatory variables. Various aspects of caliblation problem have been summarized in Brown, P.J. (1982), Brown, P.J. (1993) and Osborne, C. (1991). We assume that

$$\boldsymbol{y} = B\boldsymbol{x} + \boldsymbol{\varepsilon}, \tag{1.1}$$

where $\boldsymbol{y}$ is p dimensional random vector of response variables, B is a $p \times q$ $(p \geq q)$ matrix of unknown parameters, $\boldsymbol{x}$ is q dimensional vector of explanatory variables, and $\boldsymbol{\varepsilon}$ is a random vector of errors. Suppose that $\boldsymbol{\varepsilon}$ is distributed as p-variate normal distribution with mean $\mathbf{0}$ and covariance matrix Σ, i.e., $N_p(\mathbf{0}, \Sigma)$, where Σ is an unknown positive definite matrix. Let $Y = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n]$, $X = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n]$ $(n - q \geq p)$, and rank$X = q$, where $(\boldsymbol{y}_j, \boldsymbol{x}_j)$ $(j = 1, 2, \ldots, n)$ are observed independently. Then the model (1.1) can be written by

$$Y = BX + E\ ,$$

where $E = [\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n]$, whose columns are independent and identically distributed as $N_p(\mathbf{0}, \Sigma)$. If $\boldsymbol{y}_0$ is observed in order to estimate the corresponding explanatory variable $\boldsymbol{x}_0$, the exact confidence region has been given by Brown, P.J. (1982). The confidence region is based on the statistic

$$\frac{n-p-q+1}{p}\{1 + \boldsymbol{x}_0'(XX')^{-1}\boldsymbol{x}_0\}^{-1}(\boldsymbol{y}_0 - \hat{B}\boldsymbol{x}_0)'V^{-1}(\boldsymbol{y}_0 - \hat{B}\boldsymbol{x}_0),$$

* Department of Public Health, Miyazaki Medical College, 5200 Kihara, Kiyotake, Miyazaki, 889-1692, Japan. tel +81-985-85-0874 nakaohi@post.miyazaki-med.ac.jp

which has an F-distribution with $(p, n-p-q+1)$ degrees of freedom, where

$$\hat{B} = YX'(XX')^{-1} \quad \text{and} \quad V = Y[I_n - X'(XX')^{-1}X]Y'.$$

The statistic $\hat{B}$ is the least-squares estimate of B and is distributed as normal. The statistic V has a Wishart distribution and $S = V/(n-q)$ is an unbiased estimator of Σ (see, e.g. Brown, P.J. (1993) or Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985)). However, Brown, P.J. (1982)'s confidence region may be empty (see, e.g. Oman, S.D. (1988)). Mathew, T. and Kasala, S. (1994) have constructed an exact confidene region, which is always nonempty and is invariant under nonsingular transformation. However, their exact confidence region is not easy for practical use and may not be bounded, which is seen by a simulation study in Nakao, H. and Hyakutake, H. (1997). For approximation, Fujikoshi, Y. and Nishii, R. (1984) derived an asymptotic expansion up to the order n^{-2} for the distribution of a statistic

$$Q_1 = (\boldsymbol{t} - \boldsymbol{x}_0)'\hat{B}'S^{-1}\hat{B}(\boldsymbol{t} - \boldsymbol{x}_0), \tag{1.2}$$

where $\boldsymbol{t} = (\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}\boldsymbol{y}_0$. Nakao, H. and Hyakutake, H. (1997) have proposed an approximated confidence region based on Brown, P.J. (1982) and the statistic Q_1. Each confidence region is always an ellipsoid. Nakao, H. and Hyakutake, H. (1997) examined the accuracy of the approximations by a simulation study, in which the method obtained by Fujikoshi, Y. and Nishii, R. (1984) is sometimes under approximated. Mathew, T. and Zha, W. (1996) gave a conservative confidence region, which is nonempty.

In this paper, we assume that $\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \ldots, \boldsymbol{y}_{0l}$ are ovserved, corresponding unknown explanatory variables $\boldsymbol{x}_{01}, \boldsymbol{x}_{02}, \ldots, \boldsymbol{x}_{0l}$. Our interest is in constructing confidence regions for a linear combination $\sum c_i \boldsymbol{x}_{0i}$, where c_i's are given constants. When $l = 2, c_1 = 1, c_2 = -1$, this problem becomes that of constructing confidence regions for the difference $\boldsymbol{x}_{01} - \boldsymbol{x}_{02}$ between explanatory variables based on corresponding future observations $\boldsymbol{y}_{01}$ and $\boldsymbol{y}_{02}$. An univariate case of the set up is given in Knafl, G., Sacks, J., and Spiegelman, C. (1989). If it is assumed that $\boldsymbol{x}_{01} = \boldsymbol{x}_{02} = \cdots = \boldsymbol{x}_{0l}, c_{01} = c_{02} = \cdots = c_{0l} = 1/l$, then the l measurements $\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \ldots, \boldsymbol{y}_{0l}$ are replicated future at a single unknown explanatory variable $\boldsymbol{x}_0 (= \boldsymbol{x}_{01} = \boldsymbol{x}_{02} = \cdots = \boldsymbol{x}_{0l})$ and the confidence region is for $\boldsymbol{x}_0$. In Section 2, we derive a method for constructing approximated confidence regions of $\sum c_i \boldsymbol{x}_{0i}$. In Section 3, we give an example using *Paint Finish Data* in Brown, P.J. (1982).

2. Confidence regions

In this section, we introduce confidence regions proposed by Brown, P.J. (1982), Fujikoshi, Y. and Nishii, R. (1984)and Nakao, H. and Hyakutake, H. (1997) in case of only one future observation, at the start. Next we give an approximated confidence region for a linear combination of the explanatory variables.

$100(1-\alpha)\%$ confidence region given by Brown, P.J. (1982) is

$$\{1 + \boldsymbol{x}_0'(XX')^{-1}\boldsymbol{x}_0\}^{-1}(Q_1 + Q_0) \leq k_\alpha, \tag{2.1}$$

where $k_\alpha = (n-q)pF_{p,n-p-q+1}(\alpha)/(n-p-q+1)$, $F_{m_1,m_2}(\alpha)$ is the upper $100\alpha\%$ point of an F-distribution with (m_1, m_2) degree of freedom, and

$$Q_0 = \boldsymbol{y}_0'[S^{-1} - S^{-1}\hat{B}(\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}]\boldsymbol{y}_0.$$

For approximation to the confidence region, Fujikoshi, Y. and Nishii, R. (1984) derived an asymptotic expansion for the upper percentile point $U(\alpha)$ of a distribution of the statistic Q_1 in (1.2). They assumed that $XX'/\nu = H = O(1)$ and its limit is nonsingular with $\nu = n-q-p-1$. The asymptotic expansion for the upper $100\alpha\%$ point of Q_1 is given by

$$\begin{aligned} U(\alpha) \quad = \quad & u_\alpha + \frac{1}{2\nu}\{(4p-3q)u_\alpha + u_\alpha^2 + 2\boldsymbol{x}_0'H^{-1}\boldsymbol{x}_0u_\alpha\} \\ & + \frac{1}{24\nu^2}\left[\{24p^2 - 12(7q+4)p + 55q^2 + 36q - 4\}u_\alpha \right. \\ & \qquad \left. + (36p - 35q - 14)u_\alpha^2 + 4u_\alpha^3\right] \\ & + \frac{1}{2\nu^2}\boldsymbol{x}_0'H^{-1}\boldsymbol{x}_0\{(4p-3q)u_\alpha + u_\alpha^2\} \\ & + \frac{1}{q\nu^2}\boldsymbol{x}_0'(HB'\Sigma^{-1}BH)^{-1}\boldsymbol{x}_0(p-q)(p-q-1)u_\alpha \\ & + O(\nu^{-3}) \end{aligned} \tag{2.2}$$

where $u_\alpha = \chi_q^2(\alpha)$ is the upper 100α % point of a chi-square distribution with q degrees of freedom. Using (2.2) the confidence region is

$$Q_1 \le U(\alpha). \tag{2.3}$$

The formula (2.2) includes unknown parameters $\boldsymbol{x}_0$, B and Σ. So, the unknown parameter should be replaced by their estimates for practical use. The unknown parameters in $U(\alpha)$ are replaced by $\boldsymbol{t}$, $\hat{B}$ and S, which is denoted by $\hat{U}(\alpha)$.

An approximated confidence region given by Nakao, H. and Hyakutake, H. (1997) is

$$Q_1 \le k_\alpha\{1 + \boldsymbol{t}'(XX')^{-1}\boldsymbol{t}\} - \min\{Q_0, k_\alpha\}, \tag{2.4}$$

which are based on (2.1). It is easily seen that the confidence region (2.4) is not empty and is bounded.

Now, we wish to give an approximated confidence region for a linear combination of the explanatory variables. Under the model (1.1), we assume that $\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \dots, \boldsymbol{y}_{0l}$ are observed as future measurements and corresponding unknown explanatory variables are $\boldsymbol{x}_{01}, \boldsymbol{x}_{02}, \dots, \boldsymbol{x}_{0l}$. Our interest is to construct confidence regions for their linear combination $\sum c_i\boldsymbol{x}_{0i}$, where c_i's are given constants.

Let $Y_0 = [\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \dots, \boldsymbol{y}_{0l}]$ be a $p \times l$ matrix of future observations, $X_0 = [\boldsymbol{x}_{01}, \boldsymbol{x}_{02}, \dots, \boldsymbol{x}_{0l}]$ be a $q \times l$ matrix of unknown explanatory variables, and

$$\boldsymbol{\theta} = X_0\boldsymbol{c} = \sum_{i=1}^{l} c_i\boldsymbol{x}_{0i} \tag{2.5}$$

be a q dimensional unknown vector, where $\boldsymbol{c} = (c_1, c_2, \dots, c_l)'$ is a known l dimensional constant vector. We wish to construct the confidence region of $\boldsymbol{\theta}$.

Suppose that $\boldsymbol{y}_{0i}$ $(i = 1, 2, \ldots, l)$ are independently distributed as $N_p(B\boldsymbol{x}_{0i}, \Sigma)$ $(i = 1, 2, \ldots, l)$, then $Y_0\boldsymbol{c}$ is distributed as $N_p(BX_0\boldsymbol{c}, \boldsymbol{c}'\boldsymbol{c}\Sigma)$ (see, e.g. Theorem 1.4.4 of Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985)). Furthermore we denote

$$\begin{aligned} \boldsymbol{y}_{0\bullet} &= (\boldsymbol{c}'\boldsymbol{c})^{-\frac{1}{2}} Y_0 \boldsymbol{c} , \\ \boldsymbol{x}_{0\bullet} &= (\boldsymbol{c}'\boldsymbol{c})^{-\frac{1}{2}} X_0 \boldsymbol{c} = (\boldsymbol{c}'\boldsymbol{c})^{-\frac{1}{2}} \boldsymbol{\theta}, \end{aligned} \tag{2.6}$$

then $\boldsymbol{y}_{0\bullet}$ is distributed as $N_p(B\boldsymbol{x}_{0\bullet}, \Sigma)$. Since $(\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \cdots, \boldsymbol{y}_{0l})$ are independent of $(\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n)$, $\boldsymbol{y}_{0\bullet}$ is also independent of $(\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n)$. And once $\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \cdots, \boldsymbol{y}_{0l}$ are observed, $\boldsymbol{y}_{0\bullet}$ will be obtained. Therefore we can consider that $\boldsymbol{y}_{0\bullet}$ is observed for the new object under the model (1.1) in order to estimate the corresponding $\boldsymbol{x}_{0\bullet}$. We construct the confidence regions for $\boldsymbol{x}_{0\bullet}$, then the confidence regions for $\boldsymbol{\theta}$ can be obtained by using the transformation (2.6) because $\boldsymbol{c}$ is a known constant vector.

We can get two approximated confidence regions for $\boldsymbol{\theta}$ using each methods by Fujikoshi, Y. and Nishii, R. (1984) and Nakao, H. and Hyakutake, H. (1997). It can be cosidered that $(\boldsymbol{y}_{0\bullet}, \boldsymbol{x}_{0\bullet})$ is a future observation which satisfy the model (1.1), so that Fujikoshi, Y. and Nishii, R. (1984)'s confidence region for $\boldsymbol{x}_{0\bullet}$ is easily obtained by exchanging $(\boldsymbol{y}_0, \boldsymbol{x}_0)$ for $(\boldsymbol{y}_{0\bullet}, \boldsymbol{x}_{0\bullet})$ in the formula (2.3). It is given by

$$Q_{1\bullet} \leq \hat{U}_\bullet(\alpha) , \tag{2.7}$$

where $Q_{1\bullet} = (\boldsymbol{t}_\bullet - \boldsymbol{x}_{0\bullet})'\hat{B}'S^{-1}\hat{B}(\boldsymbol{t}_\bullet - \boldsymbol{x}_{0\bullet})$, $\boldsymbol{t}_\bullet = (\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}\boldsymbol{y}_{0\bullet}$, and

$$\begin{aligned} \hat{U}_\bullet(\alpha) &= u_\alpha + \frac{1}{2\nu}\{(4p-3q)u_\alpha + u_\alpha^2 + 2\boldsymbol{t}_\bullet' H^{-1}\boldsymbol{t}_\bullet u_\alpha\} \\ &\quad + \frac{1}{24\nu^2}\left[\{24p^2 - 12(7q+4)p + 55q^2 + 36q - 4\}u_\alpha \right. \\ &\qquad \left. +(36p - 35q - 14)u_\alpha^2 + 4u_\alpha^3\right] \\ &\quad + \frac{1}{2\nu^2}\boldsymbol{t}_\bullet' H^{-1}\boldsymbol{t}_\bullet\{(4p-3q)u_\alpha + u_\alpha^2\} \\ &\quad + \frac{1}{q\nu^2}\boldsymbol{t}_\bullet'(H\hat{B}'S^{-1}\hat{B}H)^{-1}\boldsymbol{t}_\bullet(p-q)(p-q-1)u_\alpha . \end{aligned}$$

In this asymptotic expansion $\hat{U}_\bullet(\alpha)$, H, ν and u_α are the same quantities as those of (2.2).

Also, from (2.4), the approximated confidence region for $\boldsymbol{x}_{0\bullet}$ is obtained as

$$Q_{1\bullet} \leq k_\alpha\{1 + \boldsymbol{t}_\bullet'(XX')^{-1}\boldsymbol{t}_\bullet\} - \min\{Q_{0\bullet},\ k_\alpha\} , \tag{2.8}$$

where $Q_{0\bullet} = \boldsymbol{y}_{0\bullet}'\left\{S^{-1} - S^{-1}\hat{B}(\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}\right\}\boldsymbol{y}_{0\bullet}$. We obtain the approximated confidence regions for $\boldsymbol{\theta}$ by using the transformation (2.6) for (2.7) and (2.8). The confidence regions are based on the region (2.3) by Fujikoshi, Y. and Nishii, R. (1984), and the region (2.4) by Nakao, H. and Hyakutake, H. (1997), respectively. We summarize the above discussion as the following:

Theorem 2.1 *Under the model* (1.1), *a future measurement* $Y = [\boldsymbol{y}_{01}, \boldsymbol{y}_{02}, \cdots, \boldsymbol{y}_{0l}]$ *is observed, then the approximated confidence regions* (2.3) *and* (2.4) *of* $\boldsymbol{\theta}$ *in* (2.5) *are*

$$Q_c(\boldsymbol{\theta}) \leq \hat{U}_c(\alpha) \tag{2.9}$$

and

$$Q_c(\boldsymbol{\theta}) \le k_\alpha\{\boldsymbol{c}'\boldsymbol{c} + \boldsymbol{t}_c'(XX')^{-1}\boldsymbol{t}_c\} - \min\{Q_{0c}, k_\alpha \boldsymbol{c}'\boldsymbol{c}\} , \quad (2.10)$$

respectively, where $Q_c(\boldsymbol{\theta}) = (\boldsymbol{t}_c - \boldsymbol{\theta})'\hat{B}'S^{-1}\hat{B}(\boldsymbol{t}_c - \boldsymbol{\theta})$, $\boldsymbol{t}_c = (\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}Y_0\boldsymbol{c}$, $Q_{0c} = \boldsymbol{c}'Y_0'\left\{S^{-1} - S^{-1}\hat{B}(\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}\right\}Y_0\boldsymbol{c}$, *and*

$$\begin{aligned}\hat{U}_c(\alpha) \;=\; & u_\alpha \boldsymbol{c}'\boldsymbol{c} + \frac{1}{2\nu}\{(4p-3q)u_\alpha \boldsymbol{c}'\boldsymbol{c} + u_\alpha^2 \boldsymbol{c}'\boldsymbol{c} + 2\boldsymbol{t}_c'H^{-1}\boldsymbol{t}_c u_\alpha\} \\ & + \frac{1}{24\nu^2}\left[\{24p^2 - 12(7q+4)p + 55q^2 + 36q - 4\}u_\alpha \right. \\ & \qquad \left. + (36p - 35q - 14)u_\alpha^2 + 4u_\alpha^3\right]\boldsymbol{c}'\boldsymbol{c} \\ & + \frac{1}{2\nu^2}\boldsymbol{t}_c'H^{-1}\boldsymbol{t}_c\{(4p-3q)u_\alpha + u_\alpha^2\} \\ & + \frac{1}{q\nu^2}\boldsymbol{t}_c'(H\hat{B}'S^{-1}\hat{B}H)^{-1}\boldsymbol{t}_c(p-q)(p-q-1)u_\alpha .\end{aligned}$$

3. Example *–paint finish data–*

We wish to apply the confidence regions (2.9) and (2.10) in the previous section to the *Paint Finish Data* given in Brown, P.J. (1982). Where $q = 2$ facters, pigmentation and viscosity of paint, we denote $\tilde{\boldsymbol{x}} = (\text{pigmentation}, \text{viscosity})'$, were controlled each at three levels with four replicates, and $p = 6$ responses, we denote $\tilde{\boldsymbol{y}} = (\tilde{Y}_1, \tilde{Y}_2, \tilde{Y}_3, \tilde{Y}_4, \tilde{Y}_5, \tilde{Y}_6)'$, involving optical properties were obtained. The aim in future was to be able to match the paint by taking optical measurements. The complete data has been given in Brown, P.J. (1982).

The number of observations is 36. We extract four of these observations for prediction (*Predicting Data*), whose pigmentation and viscosity are assumed to be unknown. Remaining observations are used to estimate the relationship between $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$ (*Calibrating Data*). The predicting data and the calibrating data are in Table 3.1 and Table 3.2, respectively.

We adopt the multivariate linear model with an intercept at this data following Brown, P.J. (1982). In this case, instead of (1.1), the model to be used is

$$\tilde{\boldsymbol{y}} = \boldsymbol{a} + B\tilde{\boldsymbol{x}} + \boldsymbol{\varepsilon}, \quad (3.1)$$

where $\tilde{\boldsymbol{y}}$ is a p dimensional random vector of response variables, $\tilde{\boldsymbol{x}}$ is a q dimensional vector of explanatory variables, $\boldsymbol{a}$ is a p dimensional intercept vector of unknown parameters, B and ε are the same in (1.1). Let $\tilde{Y} = [\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_{32}]$, $\tilde{X} = [\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \ldots, \tilde{\boldsymbol{x}}_{32}]$, $\tilde{Y}_0 = [\tilde{\boldsymbol{y}}_{01}, \tilde{\boldsymbol{y}}_{02}, \tilde{\boldsymbol{y}}_{03}, \tilde{\boldsymbol{y}}_{04}]$ and $\tilde{X}_0 = [\tilde{\boldsymbol{x}}_{01}, \tilde{\boldsymbol{x}}_{02}, \tilde{\boldsymbol{x}}_{03}, \tilde{\boldsymbol{x}}_{04}]$. The estimators $\hat{\boldsymbol{a}}, \hat{B}$ and S are computed as

$$\begin{aligned}\hat{\boldsymbol{a}} &= (\ 1.744,\ \ 29.622,\ \ 69.628,\ \ 37.893,\ \ 93.797,\ \ 20.563\ )' , \\ \hat{B} &= \begin{pmatrix} 0.026 & -0.262 & -1.457 & -1.247 & -3.078 & 0.428 \\ -0.062 & -4.160 & -1.760 & 0.065 & -3.718 & 1.811 \end{pmatrix}' ,\end{aligned}$$

Table 3.1. Calibrating Data

jth observation	$\tilde{x}_j$ pig	vis	$\tilde{y}_j$ $\tilde{Y}_1$	$\tilde{Y}_2$	$\tilde{Y}_3$	$\tilde{Y}_4$	$\tilde{Y}_5$	$\tilde{Y}_6$
1	−1	−1	1.88	35.0	75.0	40.94	101.0	20.0
2	−1	−1	1.88	36.1	77.0	40.60	101.0	19.8
3	−1	−1	1.87	36.8	77.0	40.57	100.5	17.5
4	−1	−1	1.87	35.3	75.8	40.68	114.0	17.5
5	−1	0	1.78	31.9	72.9	39.65	107.0	19.5
6	−1	0	1.70	29.6	72.1	39.15	93.5	21.0
7	−1	0	1.79	32.4	73.2	39.88	95.0	20.6
8	−1	1	1.63	25.7	66.7	37.22	93.5	22.0
9	−1	1	1.65	26.8	67.9	37.89	86.0	21.3
10	−1	1	1.68	27.2	67.2	38.15	84.5	19.0
11	−1	1	1.61	23.8	62.9	37.36	84.0	21.3
12	0	−1	1.79	33.5	76.0	39.09	102.0	21.0
13	0	−1	1.80	31.8	71.8	39.31	103.0	20.0
14	0	0	1.74	30.5	71.5	39.31	103.0	20.1
15	0	0	1.71	29.6	71.1	38.50	99.0	21.0
16	0	0	1.73	29.5	70.0	39.09	101.0	20.8
17	0	0	1.68	28.7	71.1	38.64	98.5	20.9
18	0	1	1.50	21.0	63.0	35.82	85.0	21.1
19	0	1	1.51	21.2	63.0	35.70	84.0	22.5
20	0	1	1.50	20.6	61.6	35.77	85.0	22.2
21	0	1	1.52	21.9	63.9	35.65	85.0	21.8
22	1	−1	1.94	35.8	74.0	38.00	101.0	19.5
23	1	−1	1.89	33.9	72.0	38.08	101.0	20.1
24	1	−1	1.92	35.0	73.0	37.93	92.5	19.5
25	1	−1	1.92	33.7	70.5	38.17	83.0	21.8
26	1	0	1.87	33.0	71.0	37.18	96.0	21.0
27	1	0	1.89	33.0	70.0	37.83	99.0	19.5
28	1	0	1.86	31.5	68.0	37.31	95.0	20.1
29	1	0	1.85	31.5	68.5	37.17	91.5	22.5
30	1	1	1.60	24.6	67.4	34.09	71.0	21.0
31	1	1	1.62	23.0	60.0	34.18	86.0	20.7
32	1	1	1.62	24.0	63.0	34.16	80.0	21.4

Table 3.2. Predicting Data

ith observation	true $\tilde{x}_{0i}$ pig	vis	$\tilde{y}_{0i}$ $\tilde{Y}_1$	$\tilde{Y}_2$	$\tilde{Y}_3$	$\tilde{Y}_4$	$\tilde{Y}_5$	$\tilde{Y}_6$
1	−1	0	1.73	30.6	72.5	39.44	93.0	18.0
2	0	−1	1.78	31.8	72.5	38.73	101.0	20.8
3	0	−1	1.77	31.3	71.8	39.12	105.0	19.8
4	1	1	1.75	27.8	64.8	34.36	82.5	20.8

and

$$S = \begin{pmatrix} 0.009 & 0.211 & 0.251 & 0.130 & 0.325 & 0.037 \\ & 5.577 & 6.103 & 2.535 & 7.120 & 0.230 \\ & & 11.692 & 4.819 & 12.930 & 1.934 \\ & & & 2.754 & 7.874 & 1.211 \\ & & & & 51.728 & 1.629 \\ & & & & & 1.698 \end{pmatrix},$$

respectively. We reduce the data to the model without the intercept by the method in Mathew, T. and Kasala, S. (1994). Let Z be a 32×31 matrix such that

$$\left(\frac{1}{\sqrt{32}}\mathbf{1}_{32}; Z\right)$$

is a 32×32 orthogonal matrix, where $\mathbf{1}_{32}$ is a 32 dimensional vector of 1's. Define $Y = \tilde{Y}Z$, $X = \tilde{X}Z$,

$$Y_0 = \left(1 + \frac{1}{32}\right)^{-\frac{1}{2}} \left(\tilde{Y}_0 - \frac{1}{32}\tilde{Y}\mathbf{1}_{32}\mathbf{1}_4'\right)$$

and

$$X_0 = \left(1 + \frac{1}{32}\right)^{-\frac{1}{2}} \left(\tilde{X}_0 - \frac{1}{32}\tilde{X}\mathbf{1}_{32}\mathbf{1}_4'\right) ,$$

then Y and Y_0 are independently distributed as $N_{6,31}(BX; \Sigma)$ and $N_{6,4}(BX_0; \Sigma)$, respectively.

Our purpose is to calibrate for four explanatory variables; $\tilde{\boldsymbol{x}}_{01}, \tilde{\boldsymbol{x}}_{02}, \tilde{\boldsymbol{x}}_{03}$ and $\tilde{\boldsymbol{x}}_{04}$ in Table 3.2, individualy, calibrate for three differences; $\tilde{\boldsymbol{x}}_{01} - \tilde{\boldsymbol{x}}_{02}, \tilde{\boldsymbol{x}}_{01} - \tilde{\boldsymbol{x}}_{03}$ and $\tilde{\boldsymbol{x}}_{04} - \tilde{\boldsymbol{x}}_{01}$, and calibrate for an average between $\tilde{\boldsymbol{x}}_{02}$ and $\tilde{\boldsymbol{x}}_{03}$. Let $\tilde{\boldsymbol{\theta}} = \tilde{X}_0\boldsymbol{c}$, then, we shall calibrate for $\tilde{\boldsymbol{\theta}}$ in eight cases when

$$\begin{aligned} \boldsymbol{c} \;=\; & (1,0,0,0)',\ (0,1,0,0)',\ (0,0,1,0)',\ (0,0,0,1)', \\ & (1,-1,0,0)',\ (1,0,-1,0)',\ (-1,0,0,1)', \\ & (0,\frac{1}{2},\frac{1}{2},0)' . \end{aligned}$$

For the reduced model we calculate the approximated confidence regions (2.9) and (2.10) for $\boldsymbol{\theta} = X_0\boldsymbol{c}$ using the method in Section 2. Since

$$\boldsymbol{\theta} = \left(1 + \frac{1}{32}\right)^{-\frac{1}{2}} \left(\tilde{\boldsymbol{\theta}} - \frac{1}{32}\tilde{X}\mathbf{1}_{32}\mathbf{1}_4'\boldsymbol{c}\right) , \tag{3.2}$$

then we get the confidence regions for $\tilde{\boldsymbol{\theta}}$ by transformation (3.2). The approximated confidence regions for $\tilde{\boldsymbol{\theta}}$ with $\alpha = 0.05$ are obtained as

$$\tilde{Q}_c(\tilde{\boldsymbol{\theta}}) \leq \left(1 + \frac{1}{32}\right) \hat{U}_c(.05) \tag{3.3}$$

and

$$\tilde{Q}_c(\tilde{\boldsymbol{\theta}}) \leq k_{.05}\{\boldsymbol{c}'\boldsymbol{c} + \boldsymbol{t}_c'(XX')^{-1}\boldsymbol{t}_c\} - \min\{Q_{0c}, k_{.05}\boldsymbol{c}'\boldsymbol{c}\} , \tag{3.4}$$

Table 3.3. The values of $\tilde{\boldsymbol{t}}_c$, $\mathcal{R}_F$ and $\mathcal{R}_N$

$\tilde{\boldsymbol{\theta}}$	true $\tilde{\boldsymbol{\theta}}$	$\tilde{\boldsymbol{t}}_c$	$\mathcal{R}_N$ in (3.6)	$\mathcal{R}_F$ in (3.5)
(Individual)				
$\tilde{\boldsymbol{x}}_{01}$	−1 0	−1.427 0.073	12.264	10.544
$\tilde{\boldsymbol{x}}_{02}$	0 −1	−0.025 −0.484	16.761	9.773
$\tilde{\boldsymbol{x}}_{03}$	0 −1	−0.346 −0.353	16.782	9.778
$\tilde{\boldsymbol{x}}_{04}$	1 1	1.912 0.352	17.725	11.322
(Difference)				
$\tilde{\boldsymbol{x}}_{01} - \tilde{\boldsymbol{x}}_{02}$	−1 1	−1.423 0.565	24.026	20.931
$\tilde{\boldsymbol{x}}_{01} - \tilde{\boldsymbol{x}}_{03}$	−1 1	−1.098 0.423	27.552	20.527
$\tilde{\boldsymbol{x}}_{04} - \tilde{\boldsymbol{x}}_{01}$	2 1	3.390 0.284	36.228	24.967
(Average)				
$\frac{\tilde{\boldsymbol{x}}_{02} + \tilde{\boldsymbol{x}}_{03}}{2}$	0 −1	−0.188 −0.394	8.383	5.084

respectively, where $\tilde{Q}_c(\tilde{\boldsymbol{\theta}}) = (\tilde{\boldsymbol{t}}_c - \tilde{\boldsymbol{\theta}})'\hat{B}'S^{-1}\hat{B}(\tilde{\boldsymbol{t}}_c - \tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{t}}_c = \left(1 + \frac{1}{32}\right)^{\frac{1}{2}} \boldsymbol{t}_c + \frac{1}{32}\tilde{X}\mathbf{1}_{32}\mathbf{1}_4\boldsymbol{c}$, $\boldsymbol{t}_c = (\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}Y_0\boldsymbol{c}$,

$$\begin{aligned}
\hat{U}_c(.05) &= u_{.05}\boldsymbol{c}'\boldsymbol{c} + \frac{1}{44}(18u_{.05}\boldsymbol{c}'\boldsymbol{c} + u_{.05}^2\boldsymbol{c}'\boldsymbol{c} + 2\boldsymbol{t}_c'H^{-1}\boldsymbol{t}_c u_{.05}) \\
&\quad + \frac{1}{2904}(61u_{.05} + 33u_{.05}^2 + u_{.05}^3)\boldsymbol{c}'\boldsymbol{c} \\
&\quad + \frac{1}{968}\boldsymbol{t}_c'H^{-1}\boldsymbol{t}_c(18u_{.05} + u_{.05}^2) \\
&\quad + \frac{3}{242}\boldsymbol{t}_c'(H\hat{B}'S^{-1}\hat{B}H)^{-1}\boldsymbol{t}_c u_{.05} \ , \\
Q_{0c} &= \boldsymbol{c}'Y_0'\left\{S^{-1} - S^{-1}\hat{B}(\hat{B}'S^{-1}\hat{B})^{-1}\hat{B}'S^{-1}\right\}Y_0\boldsymbol{c} \ ,
\end{aligned}$$

$H = 22^{-1}XX'$, $u_{.05} = \chi_2^2(.05) = 5.991$ and $k_{.05} = \frac{29}{4}F_{6,24}(.05) = 18.183$. In the resulting 95% confidence regions (3.3) and (3.4), the matrix $\hat{B}'S^{-1}\hat{B}$ is to have value

$$\hat{B}'S^{-1}\hat{B} = \begin{pmatrix} 9.827 & 5.590 \\ & 11.135 \end{pmatrix}.$$

The other values $\tilde{\boldsymbol{t}}_c, \mathcal{R}_F$ and $\mathcal{R}_N$, are in Table 3.3, where

$$\mathcal{R}_F = \left(1 + \frac{1}{32}\right)\hat{U}_c(.05) \tag{3.5}$$

and

$$\mathcal{R}_N = k_{.05}\{\boldsymbol{c}'\boldsymbol{c} + \boldsymbol{t}_c'(XX')^{-1}\boldsymbol{t}_c\} - \min\{Q_{0c}, k_{.05}\boldsymbol{c}'\boldsymbol{c}\} \tag{3.6}$$

are each of the right hand side in (3.3) and (3.4), respectively.

In all cases, both confidence regions include the true value of $\tilde{\boldsymbol{\theta}}$, and the region (3.4) include the whole of that of (3.3). However, as we could see in the simulation study given by Nakao, H. and Hyakutake, H. (1997), Fujikoshi, Y. and Nishii, R. (1984)'s confidence regions sometimes will not preserve the confidence coefficient.

Acknowledgement

The author is very grateful for various suggestion and help of Professor Hiroto Hyakutake, Kyushu University.

References

Brown, P.J. (1982). Multivariate calibration (with discussion), *Journal of Royal Statistical Society*, B**44**, 287-321.

Brown, P.J. (1993). *Measurement, Regression, and Calibration*, Oxford University Press.

Fujikoshi, Y. and Nishii, R. (1984). On the distribution of a statistic in multivariate inverse regression analysis, *Hiroshima Mathematical Journal*, **14**, 215-225.

Knafl, G., Sacks, J., and Spiegelman, C. (1989). Calibrating For Differences, in *Contributions to Probability and Statistics* (L. J. Gleser, M. D. Perlman, S. J. Press, and A. R. Sampson, eds.), 335-348.

Knafl, G., Sacks, J., Spiegelman, C., and Ylvisaker, D. (1984). Nonparametric calibration, *Technometrics*, **26**, 233-242.

Mathew, T. and Kasala, S. (1994). An exact confidence region in multivariate calibration, *The Annals of Statistics*, **22**, 94-105.

Mathew, T. and Zha, W. (1996). Conservative confidence regions in multivariate calibration, *The Annals of Statistics*, **24**, 707-725.

Nakao, H. and Hyakutake, H. (1997). Approximated confidence regions in multivariate linear calibration, *Communications in Statistics: Simulation and Computation* , **26**, 829-839.

Oman, S.D. (1988). Confidence regions in multivariate calibration, *The Annals of Statistics*, **16**, 174-187.

Osborne, C. (1991). Statistical Calibration: A Review, *International Statistical Review*, **59**, 309-336.

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

Received September 24, 2002

Revised October 4, 2002