Indian Journal of Animal Sciences **78** (4): 401–405, April 2008

Profile analysis for animal growth data

SUSHEEL KUMAR SARKAR[1] and KRISHAN LAL[2]

Indian Agricultural Statistics Research Institute, New Delhi 110 012 India

Received: 5 May 2007; Accepted: 27 November 2007

ABSTRACT

Growth of animal is a repeated measurements data and observations taken on body weight at different time points are correlated to each other. A multivariate technique called profile analysis has been discussed to analyse the growth data (body weight) of animal. This technique is simplified and the data obtained from Jabalpur and Tirupati research stations under AICRP on pigs have been analysed. It was found that the null hypothesis of no interaction of groups with time points and that there is no difference in growth between groups is accepted at 5% level of significance. But the hypothesis of no difference in growth over different time points is rejected for both the research stations. It was observed that the period of 24–32 weeks of age is the period when the growth of pigs is accelerated fast for Jabalpur research station and 12–16 weeks, 20–24 weeks and 28–32 weeks of age are the periods when the growth is accelerated fast for Tirupati research station. These are the periods that need more attention for having the better growth of pigs. Further, 2 traditional (univariate) methods of analysis of repeated measures data have been described and the merits of profile analysis over these methods have also been discussed.

Key words: Growth data, Multivariate descriptive approach, Profile analysis, Repeated measurements

The growth data on animals are the data recorded on body weight at different time points on the same animal. These data may be affected by many factors such as sex, breed *etc*. Generally multivariate techniques are used for the analysis of growth (repeated measurements) data as the observation at different time points are correlated to each other. Many approaches to the analysis of growth data have been studied by Wishart (1938), Box (1950), Rao (1958), Danford *et al.* (1960), Church (1966), Lal *et al.* (1988 a,b), who among others, have described the methods for the analysis of growth data. Most of these methods are restricted to the settings in which the response variable is normally distributed and the data are balanced and complete. The simplest approach to analyze such data is to use the technique of analysis of variance (ANOVA) separately at different time points. This approach, however, does not provide any information regarding changes in the data with time points. Another approach is to reduce the vector of multiple measurements from each experimental unit to two or three measures and the analysis is carried out using these measures.

A multivariate descriptive approach to the analysis the growth data is profile analysis (Davis 2002). In this approach, we have 3 different hypotheses. First, we test whether the different factors (groups) significantly affect the growth of animals; secondly whether there is significant difference among the time points; and thirdly whether there is any interaction between the factors and the time points.

Pigs play an important role in increasing meat (pork) production because of their high proficiency, faster growth, excellent feed convergence and shorter generation interval. Moreover, pigs are generally less prone to diseases compared to other animals and need less attention for rearing. Pigs are reared mainly for meat production, so more attention is needed for growth of pigs. In this paper, a brief description of the technique of profile analysis is given in general for animal growth data and this technique is used for analyzing the growth data on pigs maintained at Jabalpur and Tirupati locations. The time points at which the growth of pigs is accelerated are also identified. Further, the data have been analysed by using 2 traditional approaches and the merits of profile analysis have been discussed over the traditional methods.

MATERIALS AND METHODS

Data description

The data for the present study have been collected from The All-India Co-ordinated Research Project on Pigs, IVRI, Izatnagar. The data pertain to 2 research stations of AICRP on Pigs, viz. Jabalpur and Tirupati, for both male and female pigs separately. The collected data were on body weights (kg) of crossbred (Largewhite-Yorkshire × desi) pigs for both the research stations. The data were recorded weekly up to

Present address: [1]Ph.D. student, [2]Senior Scientist, Division of Design of Experiments.

pre-weaning (birth to 8 weeks of age) and 4 weekly till maturity (as follows).

Centre	Year	Male	Female	Data available up to (weeks)
Jabalpur	1987–88	25	24	36
	1988–89	17	16	20
	1989–90	6	17	36
Tirupati	1986–87	15	13	32
	1987–88	8	9	32
	1988–89	26	16	20

The data for 1988–89 of both Jabalpur and Tirupati research stations were up to 20 weeks only. These data were not to be used because the growth is insufficient to draw any valid conclusion.

Profile analysis for repeated measures data

Suppose that repeated measurements at t time points have been obtained from s groups of subjects. Let n_h be the number of subjects in-group h for $h=1,2,...,s$ and denote the total $n=\sum_{h=1}^{s} n_h$ number of animals. Let for the Y_{hij} denotes the response at time j from the ith subject in group h for h=1,2,000,s; i=1,2,...,n_h and j=1,2,...,t. We assume that the data vectors $y_h=(y_{h1},...,y_{ht})'$ are independent and normally distributed with mean $\mu_h=(\mu_{h1},...,\mu_{ht})$ and common covariance matrix Σ_{txt}. The multivariate profile model is

$$\mathbf{Y}=\mathbf{XB}+\mathbf{E} \qquad (1)$$

where, **Y** is matrix of observations with rows $y'_{11},...,y_{sns}$ and **E** is n×t matrix of errors $e'_{11},\cdots,e'_{sn_t}$ with rows and $e_{hi}\sim N_t(0,\Sigma)$, $\mathbf{X}_{n\times s}$ is design matrix and $\mathbf{B}_{s\times t}$ is matrix of unknown parameters. This model can be described as

$$\begin{pmatrix} \mathbf{y}'_{11} \\ \vdots \\ \mathbf{y}'_{1n_1} \\ \hline \mathbf{y}'_{21} \\ \vdots \\ y'_{2n_2} \\ \hline \vdots \\ \hline \mathbf{y}'_{s1} \\ \vdots \\ \mathbf{y}'_{sn_s} \end{pmatrix} = \begin{pmatrix} 1\ 0 \cdots 0 \\ \vdots\ \vdots \cdots \vdots \\ 1\ 0 \cdots 0 \\ \hline 0\ 1 \cdots 0 \\ \vdots\ \vdots \cdots \vdots \\ 0\ 1 \cdots 0 \\ \hline \vdots \\ \hline 0\ 0 \cdots 1 \\ \vdots\ \vdots \cdots \vdots \\ 0\ 0 \cdots 1 \end{pmatrix} \begin{pmatrix} \mu_{11} & \cdots & \mu_{1t} \\ \mu_{21} & \cdots & \mu_{2t} \\ \vdots & \cdots & \vdots \\ \mu_{s1} & \cdots & \mu_{st} \end{pmatrix} + \begin{pmatrix} e'_{11} \\ \vdots \\ e'_{1n_1} \\ \hline e'_{21} \\ \vdots \\ e'_{2n_2} \\ \hline \vdots \\ \hline e'_{s1} \\ \vdots \\ e'_{sn_s} \end{pmatrix} \qquad (2)$$

The profile analysis provides the detailed analysis of data by using the model (1). In this model, the 3 general null and alternative hypotheses to be tested are:

H_{01}: There is no interaction of groups with time points; H_{11}: there is interaction of groups with time points; H_{02}: there is no difference among the levels of groups; H_{12}: at least 2 levels of groups are different; H_{03}: there is no difference in growth among time points;. H_{13}: at least two time points are different in growth; and H_{01} is tested first, because testing of hypotheses H_{02} and H_{03} will depend whether H_{01} is accepted or rejected.

H_{01}: No interaction between groups by time points (Test of parallelism)

This hypothesis can be written as:

$$\mathbf{H}_{01}: \begin{pmatrix} \mu_{11}-\mu_{12} \\ \mu_{12}-\mu_{13} \\ \vdots \\ \mu_{1,t-1}-\mu_{1t} \end{pmatrix} = \begin{pmatrix} \mu_{21}-\mu_{22} \\ \mu_{22}-\mu_{23} \\ \vdots \\ \mu_{2,t-1}-\mu_{2t} \end{pmatrix} = \cdots = \begin{pmatrix} \mu_{s1}-\mu_{s2} \\ \mu_{s2}-\mu_{s3} \\ \vdots \\ \mu_{s,t-1}-\mu_{st} \end{pmatrix}. \qquad (3)$$

In terms of general hypothesis H_0: $\mathbf{ABC}\equiv\mathbf{D}$, where

$$\mathbf{A}_{(s-1)\times s}=(\mathbf{I}_{(s-1)},-\mathbf{1}_{(s-1)}),\ \mathbf{C}_{t\times(t-1)}=\begin{pmatrix} 1 & 0 \cdots 0 \\ -1 & 1 \cdots 0 \\ 0 & -1 \cdots 0 \\ \cdots & \cdots \\ 0 & 0 \cdots 1 \\ 0 & 0 \cdots -1 \end{pmatrix} \text{ and } D_{(s-1)\times(t-1)}=\begin{pmatrix} 0 \cdots 0 \\ \vdots \ddots \vdots \\ 0 \cdots 0 \end{pmatrix} \qquad (4)$$

I is identity matrix and **1** is vector of 1's.

H_{02}: Tests of no difference among groups

Depending on the results of the test of $\mathbf{H}_{01}$, two tests of hypothesis $\mathbf{H}_{02}$ of no differences among groups are possible. First, if the parallelism hypothesis is reasonable, the test for differences among groups can be carried out using the sum (or average) of the repeated observations from each subject. In this case, the null hypothesis is H_{02}: ABC ≡ D , where

$$\mathbf{A}_{(s-1)\times s}=(\mathbf{I}_{s-1},-\mathbf{1}_{s-1}),\ \mathbf{C}_{t\times 1}=\mathbf{1}_t,\ \text{and}\ \mathbf{D}_{(s-1)\times 1}=\mathbf{0}_{s-1} \qquad (5)$$

0 is vector of 0's.

Otherwise, multivariate test for differences among groups can also be carried out without assuming parallelism. In this case the null hypothesis is

$$\mathbf{H}_{02}: \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1t} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2t} \end{pmatrix} = \begin{pmatrix} \mu_{s1} \\ \mu_{s2} \\ \vdots \\ \mu_{st} \end{pmatrix} \qquad (6)$$

H_{03}: Tests of no differences among time points

Depending on the result of the test of H_{01}, two tests of

H_{03} are possible. If the parallelism hypothesis is reasonable, the test for differences among time points can be carried out using the sum (or average) across groups of the observations at each time point. In this case, the null hypothesis is H_{03}: $\mathbf{ABC} \equiv \mathbf{D}$, where

$$\mathbf{A}_1 x_s = (1,\ldots,1),\ \mathbf{C}_t x(t-1) = \begin{pmatrix} \mathbf{I}_{t-1} \\ -\mathbf{1}'_{t-1} \end{pmatrix},\ \mathbf{D}_1 \times (t-1) = \mathbf{0}'\ (1\times t-1). \quad (7)$$

The hypothesis $\mathbf{H_{03}}$ can also be tested without assuming parallelism

$$\mathbf{H}_{03}: \begin{pmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{s1} \end{pmatrix} \begin{pmatrix} \mu_{12} \\ \mu_{22} \\ \vdots \\ \mu_{s2} \end{pmatrix} \cdots = \ = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ \vdots \\ \mu_{st} \end{pmatrix}. \quad (8)$$

For testing the hypothesis $\mathbf{H}_{01}$, $\mathbf{H}_{02}$ and $\mathbf{H}_{03}$ Wilk's statistic is used.

The method described above is the general method when the data are for s groups and t time points. This method has been illustrated by using the body weight data on pigs. The data available with us are having only two levels of the group *i.e.* male and female. So the test statistic used for testing the null hypotheses $\mathbf{H}_{01}$, $\mathbf{H}_{02}$ and $\mathbf{H}_{03}$ will follow Hotelling's T^2 depending on the rank of matrix C. The degrees of freedom for the Hotelling's T^2 test statistics will vary according to three hypotheses $\mathbf{H}_{01}$, $\mathbf{H}_{02}$ and $\mathbf{H}_{03}$. Thus

$$T^2 = \frac{n_1 n_2}{n_1 + n_2}(\overline{\mathbf{Y}}_1 - \overline{\mathbf{Y}}_2)\mathbf{C}'(CSC')^{-1}C(\overline{\mathbf{Y}}_1 - \overline{\mathbf{Y}}_2) \quad (9)$$

and $$F = \frac{(n_1 + n_2 - c - 1)}{(n_1 + n_2 - 2)c} T^2 \sim F_{c, n_1 + n_2 - c - 1} \quad (10)$$

where c ($c < t$) is the rank of matrix $\mathbf{C}$, $\overline{\mathbf{Y}}_1$ and $\overline{\mathbf{Y}}_2$ are mean vectors of male and female, S is pooled estimate of covariance matrix $\mathbf{\Sigma}$ and n_1 and n_2 are numbers of male and female animals, respectively.

RESULTS AND DISCUSSION

The data collected from AICRP on Pigs, IVRI, Izatnagar for two research stations Jabalpur and Tirupati, were analysed by using profile analysis. The results obtained by this method for the 2 research stations are discussed here.

Jabalpur Research Station

The data obtained from Jabalpur research station on 10 time points i.e. 0, 4, 8, 12, 16, 20, 24, 28, 32 and 36 weeks of age were analyzed. The rank of matrix C for the null hypothesis $\mathbf{H}_{01}$, $\mathbf{H}_{02}$ and $\mathbf{H}_{03}$ from equations (4), (5) and (7). The rank of C matrix were 9, 1 and 9 respectively. The values of Hotelling's T^2 and F statistics were worked out by using equations (9) and (10) for the 3 hypotheses for the year 1987–88 and 1989–90 and are presented in Table 1. It revealed that the first hypothesis, the hypothesis of no interaction between sex and time points is accepted. This implies that the growth of the male and female is parallel over the time points. The hypothesis of no difference among groups (sex) is also accepted for both the years 1987–88 and 1989–90. But the third hypothesis of no difference in time points is rejected. We can say that body weight at different time points are varying significantly for the years 1987–88 and 1989–90.

Since the growth of pigs is significantly varying with time periods so one will be interested to know the periods when the growth is accelerated. It is because that these periods need proper care for the animals. Thus the body weight gains (difference of 2 consecutive body weights) are obtained for different time periods 0–4, 4–8, ., 32–36 weeks and the technique of analysis of covariance is used. Birth weight of animal is taken as the auxiliary variable. Also pair-wise comparisons have been made for different periods. Results of the analysis of covariance are given in Table 2a and 3 for the years 1987–88 and 1989–90, respectively. The different periods are highly significant in different years for male and female separately. From pair-wise comparisons, it is observed that the growth is accelerated at the period of 4–8 week and 28–32 for male and female pigs in 1987–88. For the year 1989–90 growth was accelerated at the period of 24–28 week for male and female both. So in general we can say that the period of 24–32 weeks is the period of accelerated growth of pigs maintained at Jabalpur research station.

Tirupati Research Station

The data from Tirupati research station were available up to 32 weeks of age of pigs so the data of 9 time points *i.e.* 0, 4, 8, 12, 16, 20, 24, 28, and 32 weeks of age, were analysed. The rank of matrix C for the null hypothesis is 8, 1 and 8 respectively. The values of Hotelling's T^2 and F statistics were worked out as discussed above. The results for the 3 hypotheses for Tirupati research station for the year 1986–

Table 1. Profile analysis of body weight of pigs at Jabalpur

Hypothesis	1987–88			1989–90		
	df	T^2 Statistics	F value	df	T^2 Statistics	F value
H_{01}	9, 39	12.202	0.180	9, 13	70.794	4.309
H_{02}	1, 47	0.223	0.986	1, 21	0.002	0.0001
H_{03}	9, 39	1251.012	101.145**	9, 13	1922.214	117.004**

Table 2. Analysis of covariance for weight gains at Jabalpur research station (1987–88)

Hypothesis	Male			Female		
	df	MS	F value	df	MS	F value
Period	8	27.74	8.63**	8	22.67	4.36**
Birth weight	1	41.45	12.89**	1	12.93	2.49
Error	215	3.22		206	5.19	

Table 3. Analysis of covariance for weight gains at Jabalpur research station (1989–90)

Hypothesis	Male			Female		
	df	MS	F value	df	MS	F value
Period	8	20.67	9.49**	8	144.20	45.10**
Birth weight	1	3.41	1.57	1	24.86	7.77**
Error	44	2.18		143	3.20	

Table 4. Profile analysis of body weight of pigs at Tirupati

Hypothesis	1986–87			1987–88		
	df	T^2 Statistics	F value	df	T^2 Statistics	F value
H_{01}	8, 19	5.408	0.416	8, 8	12.764	0.7091
H_{02}	1, 26	0.004	0.0003	1, 15	5.625	0.3125
H_{03}	8, 19	585.71	45.0546**	8, 8	1198.412	66.5784**

87 and 1987–88 are presented in the Table 4. It is apparent that the first hypothesis, which is the hypothesis of no interaction between sex and time points are accepted. We can say that the growth of male and female are parallel. The hypothesis of no difference among groups (sex) is also accepted for both the years 1986–87 and 1987–88 but the third hypothesis of no difference among time points is rejected for both the years 1986–87 and 1987–88. We can say that the body weights at different time points are varying significantly for both the years. Thus we see that results for the three hypotheses H_{01}, H_{02} and H_{03} are obtained for Tirupati research station are similar as the results obtained for Jabalpur research station. The points where the growth is accelerated are obtained by the technique discussed in previous section. Results of the analysis are given in the Table 5 and Table 6 for the years 1986–87 and 1987–88,

Table 5. Analysis of covariance for weight gains at Tirupati research station (1986–87)

Hypothesis	Male			Female		
	df	MS	F value	df	MS	F value
Period	7	68.10	9.46**	7	41.89	9.18**
Birth weight	1	70.55	9.80**	1	23.22	5.09**
Error	111	7.19		95	4.56	

Table 6. Analysis of covariance for weight gains at Tirupati research station (1987–88)

Hypothesis	Male			Female		
	df	MS	F value	df	MS	F value
Period	7	82.24	4.54**	7	59.01	2.86**
Birth weight	1	26.18	1.45	1	5.18	0.25
Error	55	18.10		63	20.60	

**indicates significant at 1% level of significance in all the Tables.

respectively. It revealed that different time periods are highly significant in different years for male and female. But if we see overall picture, the period of 12–16 week gave maximum increase in weight gain for male and female pigs for the year 1986–87 and the period of 20–24 week and 28–32 week gave the maximum increase in weight gain for male and female pigs for the year 1987–88. Thus we can say that the periods 12–16 weeks, 20–24 weeks and 28–32 weeks are the periods when the growth of pigs is accelerated and pigs need more attention.

Comparison of profile analysis with the traditional methods

Traditionally the repeated measures data are analyzed by using univariate methods because of easy understanding and calculations. These methods are simple analysis of variance (ANOVA) and reduction of vector of multiple measurements. In simple ANOVA method, analysis is carried out at each time point separately and in reduction of multiple measurements, 2 or 3 estimates of growth data are obtained. In it, quadratic regression model is fitted and the estimates of regression coefficients are obtained. The analysis is carried out on these estimates. The pig data, described above, were analysed by these 2 methods and obtained the results. A comparison is made between the profile analysis and the 2 methods and following observations were made:

1. In profile method, we have tested the significance of growth at different time points and its interaction with groups. This could not be possible by the 2 traditional (univariate) methods.
2. The analysis by simple ANOVA method dose not consider the correlation structure among the different time points while the second method considers it to some extent and the method of profile analysis considers it completely.
3. The probabilities of significance were in general smaller in the method of profile analysis than the 2 methods.

Thus in general we conclude that profile analysis is the appropriate methods for the analysis of repeated measures data.

REFERENCES

Box G E P. 1950. Problems in the analysis of growth and wear curve. *Biometrika* **6**: 362–89.

Church A J. 1966. Analysis of data when the response is a curve. *Technometrics* **8**: 229–46.

Davis C S. 2002. *Statistical Methods for the Analysis of Repeated Measurements*. Springer- Verlag, New York.

Danford M B, Haghes H M and Mcnef R C. 1960. On the analysis of repeated measurement experiments. *Biometrics,* **16**: 547–65.

Krishan Lal, Sharma, G C, Mishra R R. and Shiv Prasad 1988a. Growth rate of indigenous pigs. *Indian Journal of Animal Sciences* **58**(9): 1116–20.

Krishan Lal, Shiv Prasad, Sharma G C and Mishra R R. 1988b. Comparative study of growth rates in three generations of indigenous pigs. *Indian Journal of Animal Research* **22**(2): 85–89.

Rao C Radhakrishna 1958. Some statistical methods for comparison of growth curves. *Biometrics* **14**: 1–17.

Wishart, John 1938. Growth-rate determination in nutrition studies with the bacon pigs and their analysis. *Biometrika* **30**: 16–28.