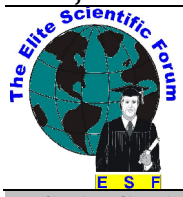
Pak. j. life soc. Sci. (2013), 11(3): 208-211 E-ISSN: 2221-7630;P-ISSN: 1727-4915



Pakistan Journal of Life and Social Sciences
www.pjlss.edu.pk

RESEARCH ARTICLE

Evaluation of Livestock Data Using Multivariate Analysis Techniques

Asghar Ali, Muhammad Aman Ullah* and Muhammad Amjad Ejaz
Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan

ARTICLE INFO
Received: Jun 26, 2013
Accepted: Nov 03, 2013
Online: Nov 08, 2013

Keywords
Average
Clusters
Correlation matrix
Principal components
Standard error

***Corresponding Author:**
aman_stat@yahoo.com

ABSTRACT

The field of livestock plays a significant role in the economy of Pakistan. The available univariate techniques are not competent to uncover all the features of the livestock data. Thus current study was conducted to assess more features of the livestock variables measured at a livestock experiment station, Rakh Gulama, Bhakkar, Pakistan which has data on 74 Holstein-Friesian x Sahiwal crossbred and 237 Nili-Ravi buffaloes for various productive and reproductive traits accumulated over 10 and 13 years respectively. The highest correlation was observed between calving interval and service period (r = 0.98), followed by calving interval and dry period (r = 0.95) and service period and dry period (r = 0.93) for different traits of buffaloes while for different traits of cattle's correlation between these variables is also high. However, the first principal component (PC_1) elucidates 41.8% of the total variation for different traits of buffalo's data while it is 44.78% for different traits of cattle's data. The PC_2, PC_3 and PC_4 for different traits of buffaloes explains 26.33%, 17.67% and 12.44% of the total variation respectively while these are 24.11%, 18.22% and 11.44% for various traits of cattle's data. On the basis of these results it was suggested that the variables under study can be classified into three clusters i.e. (i) dry period, calving interval, service period and overall average (ii) age at first calving and dry period, and (iii) milk yield, wet average and lacation period to mitigate the misleading results that are revealed using univariate techniques.

INTRODUCTION

The pivotal role of livestock in economy of every country of the world is admitted without any reservation. It provides principal source for essential items of human diet in the form milk, meat and eggs along with wool, hair, hide, skins, blood, bones as well as farmyard manure. It can easily be established from official statistics of Pakistan that more than 35 million rural population of the country is engaged in livestock having households holding 2-3 cattle/buffaloes and 5-6 sheep/goats deriving more than 40% of their income from it. The contribution of livestock is 11.6% to GDP of Pakistan, which is higher than the 10.3% contribution made by crop sector. Buffaloes are the main source for milk production in Pakistan. From the total production 4.95×10^7 tons of milk production during the year 2013, 3.05×10^7 tons was from buffaloes (GOP, 2012 and GOP, 2013).

The application of univariate statistics in the area of productivity of different breeds of buffaloes and cattle has seen a great surge of research activities during the last 20 years or so. It is clear from various articles that appeared in the literature, see for example, Hinojosa et el. (1980), Shah and Zafar (1986), Akhter et al. (1990), Shafique and Usmani (1996), Mustafa et al. (2003), Javed et al. (2006), Maji and Shaibu, 2012) and David et al. (2012).

Understanding of data from any field of life is not very straightforward. Most of the complications regarding analysis and interpretation of data can be minimized if adequate methods for the purpose are used and applied in appropriate ways. Researchers in the field of livestock mostly rely on univariate analysis, while the characteristics measured or observed are mostly correlated, hence, due consideration is required to be given to such a vital factor, otherwise faulty results can occur and the inference, therefore, will be misleading. Therefore, the objective of this study was to suggest most appropriate techniques for the analysis of livestock data to mitigate the misleading results that are revealed using univariate techniques.

MATERIALS AND METHODS

Data description

The data on 74 Holstein-Friesian x Sahiwal crossbred and 237 Nili-Ravi buffaloes for various productive and reproductive traits accumulated over 10 and 13 years respectively observed at Livestock Experiment station, Rakh Gulama, Bhakkar was available to us.

Statistical analyses

The data on 74 Holstein-Friesian x Sahiwal crossbred and 237 Nili-Ravi buffaloes for various productive and reproductive traits were subjected to statistical analyses, using multivariate techniques. The principal components PC_1, PC_2, PC_3 and PC_4 and correlation matrix were computed (Mardia et al., 1979; Ali et al, 1985; Jolliffe, 2002; Ejaz, 2005; Maji and Shaibu, 2012). The statistical software MINITAB was used for the analysis of data.

RESULTS AND DISCUSSION

Using the approach of the researchers in the field of livestock, Means ± SD of such data are presented in Tables 1 and 2. Going through these tables we find similar interpretation of the results as have been observed by the various researchers in their studies (Mardia et al., 1979; Ali et al., 1985; Jolliffe, 2002). To our point of view, there are more information available than that can be explained from the Tables 1 and 2; such information are available in tables of correlation matrices (Tables 3a, 3b) and in the results of principal component analysis (Table 4).

Now, going though the results had shown in the correlation matrices (Tables 3a and 3b). These tables show linear relationship among the multivariate variables of interest. It is very straight to observe from both of the tables that variables 'dry period', 'calving interval' and 'service period' are highly positively correlated among each other whereas all these three variables are negatively correlated with 'overall average'. Another group of variables is 'age at first calving' and 'parity number' has very strong and positive relationship between each other while these variables have tendency of negative relationship with the earlier mentioned variables and have no relationship with the remaining set of variables. The remaining set of variables, which includes 'milk yield', 'wet average' and 'lactation period' has such a correlation structure that 'milk yield' positively and moderately high correlation with both of the other variables i.e. 'wet average' and 'lactation period' while these two are uncorrelated with each other.

Table 4 presents the results of Principal components where we obtained that the first PC for different traits of buffaloes explains 41.8% of the total variation in the data. It has high loadings for 'dry period' (+), 'calving interval' (+), 'service period' (+) and 'overall average' (-); hence justify the finding from the correlation matrix as stated above. The remaining variables other than 'lactation period' are with the same (-) signs but with low loadings while the variable 'lactation period' has the smallest loading in this PC.

Second PC which was interpreted as the general index of the data explains about 26.3% of the variation as all the variables are with the same size and the magnitude of the coefficients of this component varies from 0.19 to 0.50 (reasonably high) hence are all important to be considered as the representative of the second PC. Third PC is a comparison of 'parity number' and 'age at first calving' with 'overall average', 'milk yield', and 'water average' and 'lactation period'. This PC explains 17.7% of the total variation; hence is considerably important to take into account. If we go through the coefficients of fourth PC, which explains about 12.4% of the total variation, it says that there is a contrast of 'lactation period' with 'water average' and 'dry period'. The remaining PCs, which altogether explained less than 02% of the total variations, are therefore unimportant and have nothing to add in the interpretation of the results.

The results for cattle data as shown in Table 4 were almost the same with minor difference with the above-explained results for buffaloes' data. It was observed that both for buffaloes and cattle data were with the same correlation structure. If all these results and interpretations are pooled up, reach to the following findings:

The variables under study can be classified into three clusters i.e. (i) 'dry period', 'calving interval', 'service

Table 1: Means (±S.E) of different traits of buffaloes in different parities

Variables	Parity Number (Number of Buffaloes)					
	1 (74)	2 (66)	3 (61)	4 (45)	5(27)	6(12)
Lactation Period (Days)	253.1±10.3	284.53±8.98	292.39±7.28	286.51±7.36	282.19±9.43	306.5±18.6
Milk Yield (Liters)	1396.2±70.42	1728.6±87.3	1890.8±69.8	1959.4±91.9	1978±144	1766±130
Dry Period (Days)	314.2±20.84	239.2±15.9	221.3±13.8	195.5±18.9	196.8±19.3	104.1±18.2
Service Period (Days)	287.21±16.55	244±14.9	228.1±16.0	194.6±21.1	199.2±18.2	124.3±25.2
Calving Interval (Days)	567.3±16.47	523.7±15	513.7±15.9	482±20.3	479.10±17.9	410.10±17.9
Wet Average (Liters)	5.51±0.15	6.10±0.2	6.47±0.2	6.86±0.27	6.92±0.389	5.80±0.33
Overall Average (Liters)	2.68±0.17	3.43±0.17	3.89±.21	4.31±0.25	4.215±0.32	4.41±0.36
Ag e at 1st Calving (Months)	45.13±1.29	62.36±1.50	76.39±1.32	92.75±1.74	104.96±21.96	117.16±3.83

Table 2: Means (±S.E) of different traits of cattle in different parities

Variables	Parity Number (Number of Buffaloes)							
	1 (147)	2 (171)	3 (149)	4 (97)	5(67)	6(46)	7(27)	8(19)
Lactation Period (Days)	290.75 ±5.43	307.4 ±4.69	302.72 ±4.24	296.58 ±6.2	294.58 ±6.2	299.91 ±8.4	296.74 ±8.2	291.12 ±42.3
Milk Yield (Liters)	1313.52 ±31.88	1480.31 ±34.49	1538.72 ±34.14	1516.14 ±36.62	1592.05 ±59.48	1551.09 ±60.19	1710.01 ±86.86	1623 ±74.56
Dry Period (Days)	425.53 ±15.32	354.42 ±11.11	348.33 ±14.93	616.20 ±14.39	302.32 ±17.56	278.93 ±24.73	278.53 ±31.04	253.16 ±44.1
Service Period (Days)	696.87 ±4.88	351.28 ±11.86	337.72 ±15.07	299.44 ±14.25	285.65 ±19.16	268.86 ±26.68	260.03 ±34.42	236.16 ±45.27
Calving Interval (Days)	716.3 ±14.88	661.79 ±11.94	651.05 ±15.02	612.9 ±14.28	596.91 ±20.78	578.3 ±19.07	575.29 ±34.47	544.73 ±27.52
Wet Average (Liters)	4.56 ±0.09	4.82 ±0.09	5.11 ±0.10	5.14 ±0.11	5.46 ±0.20	5.22 ±0.18	5.77 ±0.25	05.010 ±0.04
Overall Average (Liters)	1.93 ±0.05	2.32 ±0.06	2.53 ±0.07	2.58 ±0.08	2.81 ±0.12	2.88 ±0.14	3.16 ±0.21	3.23 ±0.22
Ag e at 1st Calving (Months)	58.82 ±0.82	80.6 ±0.87	98.77 ±1.05	117.60 ±1.52	130.13 ±1.59	144.80 ±1.84	157.14 ±1.80	170.21 ±1.71

Table 3a: Correlation matrix between different traits of buffaloes

Variables	LP	MY	DP	SP	CI	WA	OA	AAC	P.No
Lactation Period	1	0.52	-0.03	0.28	0.29	-0.14	0.10	-0.01	0.01
Milk Yield	0.52	1	-0.04	0.13	-0.12	0.76	0.64	0.21	0.20
Dry Period	-0.03	-0.04	1	0.93	0.95	-0.03	-0.63	-0.23	-0.26
Service Period	0.28	0.13	0.93	1	0.98	-0.06	-0.58	-0.21	-0.25
Calving Interval	0.29	-0.12	0.95	0.98	1	-0.07	-0.60	-0.21	0.25
Wet Average	-0.14	0.76	-0.03	-0.06	-0.07	1	0.67	0.23	0.23
Overall Average	0.10	0.64	-0.63	-0.58	-0.60	0.67	1	0.33	0.37
Ag e at 1st Calving	-0.01	0.21	-0.23	-0.21	-0.21	0.23	0.33	1	0.92
Parity Number	0.01	0.20	-0.26	-0.25	0.25	0.23	0.37	0.92	1

Table 3b: Correlation matrix between different traits of cattle

Variables	LP	MY	DP	SP	CI	WA	OA	AAC	P.No.
Lactation Period	1	0.68	-0.42	0.08	0.08	0.07	0.45	0.15	0.16
Milk Yield	0.68	1	-0.38	0.06	-0.05	0.76	0.83	0.13	0.21
Dry Period	-0.42	-0.38	1	0.87	0.88	-0.14	-0.70	-0.22	-0.33
Service Period	0.08	0.06	0.87	1	0.99	-0.13	-0.53	-0.17	-0.29
Calving Interval	0.08	-0.05	0.88	0.99	1	-0.12	-0.53	-0.16	-0.28
Wet Average	0.07	0.76	-0.14	-0.13	-0.12	1	0.73	0.06	0.16
Overall Average	0.45	0.83	-0.70	-0.53	-0.53	0.73	1	0.17	0.30
Ag e at 1st Calving	0.15	0.13	-0.22	-0.17	-0.16	0.06	0.17	1	0.90
Parity Number	0.16	0.21	-0.33	-0.29	-0.28	0.16	0.30	0.90	1

Table 4: Principal Component Analysis for Buffaloes and Cattle data sets

Variables	Principal Components (Buffaloes)				Principal Components (Cattle)			
	PC_1	PC_2	PC_3	PC_4	PC_1	PC_2	PC_3	PC_4
Lactation Period	0.042	0.301	-0.239	-0.787	-0.211	-0.374	-0.061	-0.697
Milk Yield	-0.203	0.502	-0.377	-0.049	-0.318	-0.520	0.070	-0.0004
Dry Period	0.447	0.250	0.241	0.283	0.442	-0.200	-0.133	0.313
Service Period	0.438	0.335	-0.054	0.029	0.364	-0.443	-0.164	-0.045
Calving Interval	0.427	0.350	0.201	-0.004	0.372	-0.419	-0.177	-0.033
Wet Average	-0.234	0.340	-0.228	0.533	-0.262	-0.361	0.150	0. 617
Overall Average	-0.452	0.188	-0.201	0.051	-0.451	-0.205	0.173	0.109
Ag e at 1st Calving	-0.268	0.254	0.548	-0.056	-0.207	0.050	-0.688	0.073
Parity Number	-0.227	0.319	0.604	-0.080	-0.270	0.058	-0.624	0.120
Eigen values	3.76	2.37	1.59	1.12	4.03	2.17	1.64	1.03
%	41.78	26.33	17.67	12.44	44.78	24.11	18.22	11.44

period' and 'overall average'; note that 'overall average' will be negatively and seriously affected if there is no balance among 'dry period', 'calving interval' and 'service period'; (ii) 'age at first calving' and 'parity number'; (iii) 'milk yield', 'wet average' and 'lactation period'.

The results from PC's revealed that it was not just enough to take into account all the factors involved at equal level; some factors were supposed to be taken at priority basis and in presence of significant influence of one factor, other factors were needed to be adjusted accordingly.

Conclusions

In this article, it has been shown that multivariate data under consideration can be classified into some useful cluster using correlation matrix through principal component analysis. The variables 'dry period', 'calving interval' and 'service period' are highly correlated with the other variables, and hence should be treated alike. It is further concluded, on the basis of negative relationship, overall average of the subject under study may suffer if there is no balance among dry, calving and service periods. Finally, milk yield is strongly based on wet average and the lactation period.

REFERENCES

Akhter P, MA Khan and G Mohy-ud-Din, 1990. Heritability of Some Reproductive Traits in Nili-Ravi Buffaloes. Pakistan Veterinary Journal, 10:132-134.

Ali A, GM Clarke and K Trustum, 1985. Principal Component Analysis and Applied to Some Data from Fruit Nutrition Experiments. Statistician, 34: 365-369.

David JA, AJ Dougill and LC Stringer, 2012. Using Principal Component Analysis for information-rich socio-ecological vulnerability mapping in Southern Africa. Applied Geography, 35: 515-524.

Ejaz MA, 2005. Evaluation of Livestock Data using Multivariate Analysis Techniques. M. Phil Thesis, Bahauddin Zakariya University Multan, Pakistan.

GOP, 2012. Agriculture Census Organization, 2011-2012. Statistics Division, Government of Pakistan, Lahore.

GOP, 2013. Economic Survey of Pakistan, 2012-2013. Economic Affairs Division, Government of Pakistan, Islamabad.

Hinojosa A, A Franco and I Bolio, 1980. Genetic and Environmental Factors Affecting Calving Interval. Tropical Animal Production, 5: 165-168.

Javed MT, M Usman, M Irfan and M Cagiola, 2006. A study on tuberculosis in buffaloes: Some epidemiological aspects, along with haematological and serum protein changes. Veterinary Archive, 76: 193-206.

Jolliffe IT, 2002. Principal Component Analysis. Springer, New York, USA.

Maji AT and AA Shaibu, 2012. Application of principal component analysis for rice germplasm characterization and evaluation. Journal of Plant Breeding and Crop Science, 4: 87-93.

Mardia KV, JT Kent and JM Bibby, 1979. Multivariate Analysis. Academic Press London, UK.

Mustafa MI, M Latif, MK Bashir and Ahmed, 2003. Reproductive Performance of Red Sindhi Cattle under Hot and Humid Environment of Balochistan Province of Pakistan, Pakistan Veterinary Journal, 23: 66-72.

Shafique P and RH Usmani, 1996. Reproductive Performance and Production Nili-Ravi Buffaloes Maintained under Barani (Rain Fed). Pakistan Veterinary Journal, 16: 168-171.

Shah IH and AH Zafar, 1986. Inheritance of Age at First Calving and First Lactation Yield Sahiwal Cows. Pakistan Veterinary Journal, 6: 60-62.