A DISCRIMINANT FUNCTION FOR QUALIFICATION OF ZEBU SEMEN SAMPLES BEFORE FREEZING[1]

JOÃO FLORIANO CASAGRANDE[2], ROBERTO GOMES DA SILVA[3] and ANTONIO MIES FILHO[4]

ABSTRACT.- A discriminant function is presented for the selection of semen ejaculates from Zebu bulls before dilution and freezing. The estimated function was $\underset{\sim}{L} = 100 + 0.248\, X_1 - 0.237\, X_2 - 0.142\, X_3$, where X_1 is the percent initial motility, X_2 the proportion of primary abnormalities, and X_3 the proportion of secondary abnormalities, these variables being previously submitted to an arc-sin transformation. The above function was tested in terms of the probability of misclassification of any ejaculate and is considered suitable for practical use.

Index terms: Discriminant function, qualification, zebu semen.

INTRODUCTION

The production of frozen Zebu semen is a business of increasing importance in Brazil as the use of artificial insemination increases. One of the main remaining problems is the frequent loss of semen doses discarded because of poor quality. Except for cases of careless handling, the main cause of such losses lies in the intrinsic quality of the sperm to be frozen.

Unfortunately, little attention has been paid to the importance of the quality of an ejaculate before its dilution and freezing. However Hahn *et al.* (1969), have established a predictive equation for the proportion of motile spermatozoa after freezing. Their equation is useful for the Holstein breed only, and its efficiency is not very great ($R^2 = .54$).

Casagrande (1973) recorded 12 morphological and physical characteristics on 1,000 ejaculates of Zebu bulls established that the more important of these (i.e., the ones more linked to the survival rate of spermatozoa after freezing) were the initial motility, and the proportions of primary and secondary abnormalities. The correlations of these traits with the freezability were .648, −.508, and −.660 respectively. Consequently he suggested that these three traits could be used to judge the quality of Zebu semen before deep freezing.

In this paper these traits are used to establish a criterion for judging semen ejaculates taken from Zebu bulls bred under the usual management conditions in the State of São Paulo, Brazil.

MATERIAL AND METHODS

The data used were those recorded by Casagrande (1973) on 1,000 ejaculates from 134 bulls of the following breeds: Gir, Guzerath, Nelore, Red Sindhi, Indubrasil, and Tabapuan. These bulls were normally used for semen production at a commercial stud in the State of São Paulo. Details about the techniques used for collection and preparation of the semen samples may be obtained from one of the authors (J.F.C.) on request.

The semen samples were carefully examined for 12 characteristics both before and after 24 hours of freezing in liquid nitrogen (−196°C). They were classified as "rejected" or "approved" after freezing, depending on whether they had more or less than 20 x 10^6 motile spermatozoa, and/or progressive motility above or below 20 per cent. Since all measurements were percentages, an angular transformation of the original data was carried out before statistical analyses were performed. The multivariate techniques described by Rao (1952) and Seal (1966) were used. To help the reader un familiar with these techniques, a detailed explanation is given below.

Considering the following characteristics of the ejaculates:

X_1 – initial progressive motility,

X_2 – proportion of primary abnormalities,

X_3 – proportion of secondary abnormalities, and

Y – freezability,

the present problem is to carry on a comparison between the evarages of the two groups of ejaculates, i.e. the approved and the rejected.

Multivariate comparisons between h groups are usually subjected to the assumption that the variances and covariances are homogeneous, i.e., that

$$\underset{\sim}{S}^{(1)} = \underset{\sim}{S}^{(2)} = \ldots = \underset{\sim}{S}^{(h)},$$

where $\underset{\sim}{S}^{(i)}$ is the p x p variance-covariance matrix of the i^{th} group ($i = 1, \ldots, h$), and p is the number of considered traits. The test for this hypothesis of equality of the variance-covariance matrices in the special case of $h = 2$ is

$$X^2 = -2k \log_e W,$$

with $(p^2 + p)/2$ degrees of freedom and where

$$k = 1 - \sum_{i=1}^{h} (n_i - 1)^{-1} (N-2)^{-1} \cdot \frac{2p^2 - 3p - 1}{6(p+1)},$$

$$w = \prod_{i=1}^{h} \left[|\underset{\sim}{S}^{(i)}|^{(n_i - 1)/2} \right] / |\underset{\sim}{S}|^{(N-2)/2},$$

$$\underset{\sim}{S} = (N-2)^{-1} \sum_{i=1}^{h} (n_i - 1)\, \underset{\sim}{S}^{(i)},$$

and N is the total number of data ($N = n_1 + n_2$).

[1] Aceito para publicação em 15 de julho de 1976.

[2] Veterinarian, Semen do Brasil S.A. (SEMBRA), 14780 Barretos, São Paulo.

[3] Veterinarian, Assistant Professor, Department of Animal Science, Faculdade de Medicina Veterinária e Agronomia, Cx. Postal 145, Jaboticabal, São Paulo.

[4] Professor of Physiology and Pathology of Reproduction, Faculdade de Medicina Veterinária, Universidade Federal do Rio Grande do Sul, Cx. Postal 776, 90000 Porto Alegre, Rio Grande do Sul.

If the matrices are determined as homogeneous (cases in which X^2 is not significant), then we can perform the Hotelling's test to test the hypothesis of equality of the group means:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} \underset{\sim}{d}' \underset{\sim}{S}^{-1} \underset{\sim}{d},$$

where $\underset{\sim}{d} = \underset{\sim}{m}^{(1)} - \underset{\sim}{m}^{(2)}$, being $\underset{\sim}{m}^{(i)}$ the vector of the means of the i^{th} group. The above value is to be compared with

$$T^2_a = \frac{(N-2)p(F_{p;N-p-1})}{N-p-1},$$

where a is the significance level (0.01 or 0.05), and $F_{p;N-p-1}$ is the tabled F value at p and $(N-p-1)$ degrees of freedom. It is interesting to note that

$$F_{p;N-p-1} = \frac{N-P-1}{p(N-2)} \quad T^2.$$

The significance of the T^2 value is an indication of the existence of significant differences beween the group means.

The separation of any ejaculate into an "approved" or "rejected" group can be made using a discriminant function which gives the maximum separation between the groups. According to Fisher (1936), such a function is

$$\underset{\sim}{L} = \underset{\sim}{S}^{-1}\left[\underset{\sim}{m}^{(1)} - \underset{\sim}{m}^{(2)}\right] = \underset{\sim}{S}^{-1} \underset{\sim}{d}.$$

If we assume a given selection threshold, i.e., a value of $\underset{\sim}{L}$ bordering both groups, any random sample of semen can be rejected or approved for freezing, depending upon the respective value of $\underset{\sim}{L}$ being lower or higher than the threshold value, respectively.

To be useful, the function must be significant. This can be tested (Rao 1952) by

$$F' = \frac{1-v}{y} \cdot \frac{m+2z}{2r},$$

where

$r = p/2,$

$z = (2-p)/4$

$m = N-1-\frac{p+2}{2}$

$y =$ Wilk's criterion $= |\underset{\sim}{S}^d| \, / \, |\underset{\sim}{S}^e + \underset{\sim}{S}^d|$

$\underset{\sim}{S}^e$ = "between" groups variance-covariance matrix,

$\underset{\sim}{S}^d$ = "within" groups variance-covariance matrix.

The value of F' is compared to a tabulated F at $2r$ and $(m + 2z)$ degrees of freedom.

Results and Discussion

The average results of the "approved" and rejected" samples are presented in Table 1. It can be seen that the group averages were very different; the rejected samples showed lower motility, and higher proportions of primary and secondary abnormalities than the approved ones.

The homogeneity test showed that

$X^2 = 1.932$, with 6 degrees of freedom.

This is not significant, hence the hypothesis $\underset{\sim}{S}^{(1)} = \underset{\sim}{S}^{(2)}$ cannot be rejected. The matrices $\underset{\sim}{S}^{(1)}$, $\underset{\sim}{S}^{(2)}$ and $\underset{\sim}{S}$ are show in tables 2, 3 and 4 respectively.

TABLE 1. *Characteristics of 1,000 ejaculates from 134 Zebu bulls used in artificial insemination. Averages (above) and standard errors (below) in the original units (percent)*

Sample type	Number of samples	Motility (%)	Primary abnormalities (%)	Secondary abnormalities (%)	Freezability (%)
Approved	718	67.002	4.020	4.303	58.337
		0.001	0.001	0.009	0.004
Rejected	282	47.203	8.004	13.523	21.730
		0.008	0.003	0.009	0.011

TABLE 2. *Variance-covariance matrix* $\underset{\sim}{S}^{(1)}$ *of the approved semen samples*

Traits	X_1	X_2	X_3
X_1: motility	21.4268		
X_2: primary abnormalities	—2.9692	11.6542	
X_3: secondary abnormalities	—5.8331	2.5057	21.3850

TABLE 3. *Variance-covariance matrix* $\underset{\sim}{S}^{(2)}$ *of the rejected semen samples*

Traits	X_1	X_2	X_3
X_1: motility	70.9688		
X_2: primary abnormalities	— 1.4606	30.8721	
X_3: secondary abnormalities	—39.6968	—1.4043	82.2542

TABLE 4. *Common variance-covariance matrix $\underset{\sim}{S}$ of the 1,000 semen samples*

Traits	X_1	X_2	X_3
X_1: motility	35.3760		
X_2: primary abnormalities	— 2.5444	17.0653	
X_3: secondary abnormalities	—15.3678	1.4048	38.5235

Therefore, the Hotelling's test was performed using the common matrix $\underset{\sim}{S}$, and this gave a value for

$$T = 33.002.$$

This is highly significant ($P < 0.05$), indicating that the differences between the mean vectors of both semen groups were highly significant, and thus the three selected characteristics (motility, and proportions of primary and secondary abnormalities) can be used to establish the suitability of any semen sample for freezing.

We obtained the following discriminant function:

$$\underset{\sim}{L} = 100 + 0.248\ X_1 - 0.237\ X_2 - 0.142\ X_3,$$

which was highly significant ($P << 0.01$), indicating that the above function gives a good separation between both groups of ejeculates.

To test the practical usefulness of the estimated function, we applied it to 1,000 semen ejaculates samples from the same, population as the original data. These samples were classified as "rejected" or "approved" by direct determination of their characteristics after freezing. The respective values were then plotted on a graph (Fig. 1). This showed that the values obtained agree very well with corresponding theoretical normal values.

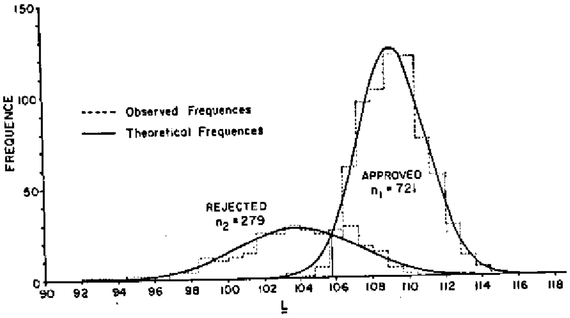

FIG. 1. *Distribution of the values of the discriminant function $\underset{\sim}{L}$ in a sample of 1,000 ejaculates previously classified into "approved" and "rejected" groups.*

Fig. 1 also shows that the distributions of both groups meet, at $\underset{\sim}{L} = 105.75$. Thus, this value can be considered a selection threshold, i.e. semen samples should be accepted for freezing if $\underset{\sim}{L} > 105.75$. The superimposition of the distributions, however, show that the discrimination is not perfect and that some probability of misclassification exists.

In fact, if P_1 is the estimated proportion of "approved" type semen ejaculates in the population, and P_2 is the proportion of "rejected" ones then assuming that our population samples is representative,

$$P_1 = 0.718$$

$$P_2 = 0.282.$$

The probability of misclassification of a sample belonging to one of the groups can be estimated by means of the graph of Fig. 1 and the table of the areas of normal curve. It was found that the probabilities of wrong classification for a sample belonging to the "approved" type and for one belonging to the "rejected" type were respectively:

$$a_1 = 0.0638$$

$$a_2 = 0.2900$$

Therefore, the probability of misclassification of any random samples is

$$P_1a_1 + P_2a_2 = 0.128,$$

which is relatively high.

Consequently, a technician would be able to accept or reject semen samples for freezing based on the threshold value of $\underset{\sim}{L} = 105.75$, but he would discover eventually that from 100 ejaculates he accepted for freezing, about 8 of them would be of poor quality; on the other hand, from the 100 samples he rejected, about 5 would be of high quality. Such errors could be avaided if *two* threshold values are considered:

a) unconditional rejection of the sample when $L \leqslant 104.0$, and

b) unconditional approval of the sample when $L \geqslant 110.0$.

Within these limits the approval or rejection of a sample would be conditional to the probability of the sample belonging to the "approved" type. This probability is equal to the opposite one at the level of $L = 105.75$, but increases in direct proportion to the $\underset{\sim}{L}$ values. A summarized table is presented for the probabilities lying in the region from $\underset{\sim}{L} = 104.0$ to $\underset{\sim}{L} = 110.0$ (Table 5).

TABLE 5. *Probabilities of any random samples of Zebu semen belonging to an approved group, P(A), given its $\underset{\sim}{L}$ value*

$\underset{\sim}{L}$	P(A)	$\underset{\sim}{L}$	P(A)
104.0	0.10	107.0	0.76
104.5	0.16	107.5	0.85
105.0	0.28	108.0	0.90
105.5	0.43	108.5	0.92
105.75	0.50	109.0	0.95
106.0	0.56	109.5	0.95
106.5	0.70	110.0	0.96

Finally, it must be remembered that the discriminant function presented above can be applied only to data that has been angularty transformed. Such transformation can be obtained from various reference tables (Snedecor 1964, Fisher & Yates 1971).

ACKNOWLEDGEMENTS

We are indebted to Dr. W.E. Kerr and Dr. F.A.M. Duarte for permission to use the computing facilities of the Department of Genetics, Faculdade de Medicina, Ribeirão Preto, SP; to the Data Processing Center of Escola de Engenharia at São Carlos; to Fundação de Amparo à Pesquisa do Estado de São Paulo, and finally to Dr. G. Siekmann, who revised the manuscript.

REFERENCES

Casagrande J.F. 1973. Relações entre algumas características físicas e morfológicas do sêmen de zebuínos e sua congelabilidade. Unpublished D. Sci. thesis, Fac. Med. Vet. Agron. Jaboticabal, SP.

Fisher R.A. 1936. The use of multiple measurement in taxonomic problems. Ann. Eugen. 7:179-188.

Fisher R.A. & Yates F. 1971. Tabelas estatísticas para pesquisas em biologia, medicina e agricultura. Brasilian translation, Ed. Univ. São Paulo-Ed. Polígono, São Paulo, p. 78-79.

Hahn J., Foote R.H. & Seidel G.E. 1969. Quality and freezability of semen from growing and aged dairy bulls. J. Dairy Sci. 52:1843-1848.

Rao C.R. 1952. Advanced statistical methods in biometric research. J. Wiley, New York, p. 236-257.

Seal H. 1966. Multivariate statistical analysis for biologists. Methuen, London.

Snedecor G.W. 1964. Metodos estatísticos. Spanish translation, Ed. Continental, Mexico, p. 377-378.

SINOPSE.- Casagrande, J.F.; Silva, R.G.da; Mies Filho, A. [*Uma função discriminante para qualificação de semen de touros zebuínos antes do congelamento*]. A discriminant function for qualification of zebu semen samples before freezing. *Pesquisa Agropecuária Brasileira, Série Zootecnia* (1976) *11*, 25-28 [En, pt] Fac. Medicina Veterinária e Agronomia, Cx. Postal 145, 14870 Jaboticabal, SP, Brasil.

É apresentada uma função discriminante para qualificação e seleção de ejaculados de sêmen de touros zebuínos antes de sua diluição e congelação, visando economia de tempo e de material. A função estimada é $\underset{\sim}{L} = 0{,}248\ X_1 - 0{,}237\ X_2 - 0{,}142\ X_3$, onde X_1 representa a percentagem de motilidade progressiva inicial, X_2 a percentagem de anomalias primárias e X_3 a proporção de anomalias secundárias dos espermatozóides antes da congelação, devendo estes dados ser previamente transformados em ângulos de arco-seno. A função acima foi testada em termos de probabilidade de erro de classificação de uma amostra qualquer de sêmen em "aproveitável" ou "não aproveitável" para congelação, tendo sido verificado ser a mesma um bom critério de seleção, desde que sejam usados dois valores-limites: se $\underset{\sim}{L} \leq 104{,}0$, a amostra deve ser rejeitada, e se $\underset{\sim}{L} \geq 110{,}0$ a amostra deve ser aprovada para fins de congelação; valores de $\underset{\sim}{L}$ entre esses limites levam à necessidade de recorrer à tabela de probabilidades, fornecida no trabalho, para se decidir quanto ao destino do ejaculado.

Termos de indexação: Função discriminante, qualificação, sêmen de zebu.