Australian Jour nal of Basic and Applied Sciences
2018 April; 12(4): pages 19-21
DOI: 10.22587/ajbas.2018.12.4.4
Research Article AENSI Publications

Discriminant Function Using Fisher's Approach For Three Multivariate Normal Populations

[1]Khawla Mustafa Sadek and [2]BasimaIdris Ismail

[1,2]*Department of Mathematical Education College Mosul University, Iraq*

Correspondence Author: Khawla Mustafa Sadek, Department of Mathematical Education College Mosul University, Iraq

Received date: 23 February 2018, **Accepted date:** 1 April 2018, **Online date:** 15 April 2018

Abstract

Background: Discriminant analysis is used in situations where the populations or groups are known apriori. one of the aims of discriminant analysis is to classify an observation, or several observations, in to these known populations, In this case we know that it or they have come from exatly one of those populations but we do not know from which. The other aim is to interpret the differences between the populations, in terms of a few linear functions of the orginal variables. In the following we use classification instead of discrimination although kendall (1966) has extensive notes the terminology.

Key words: Multivarate Normal distributiong, Baye'sTheorm, DiscriminatFunction, Fisher's Approach.

Abstract:

We consider a classification of q groupe. Let y ∈ {1, 2,…,q} be adependent variable associated with each object or observation.

Suppose that the group conditional population densities f_i (x)= f (x/y =i) and the population prior ρ_i are unknown. The Baye'stheorm may be used to express the posterior probability of population i given observation vector x.

f (i/x)=$\frac{\rho_i f_i(x)}{\sum_{\ell=1}^{q}\rho_\ell f_\ell(x)}$

If the distribution of x is continuous, it is sufficient to replace x_o by a δ neighborhood No (x_o) of $x_o$and to tend δ→ o. Baye's rule predicts the population of an observation vector x by that with the highest posterior probability max $\rho_i f_i$ (x)= $\rho_i f_i$ (x) ⇒$x \in G_i$.

Fisher's Approach when the π_i distributed multivariate normal have the same covariance matrix ∑, the (sample) linear discriminant Function's $\grave{b}_i$ X, i= 1,…, m = min{p, q-1} were introduced by fisher (1936) achieve a good separation such functions were found by condsidering the linear combination $\grave{b}$x that maximizes the ratio of the between- group sum of squares to the within- group sum of squares.

Objective:

The objective of our research is when we Let

$w_{ij}(x) = (\bar{x}_i - \bar{x}_j)\Sigma^{-1}\left[x - \frac{1}{2}(\bar{x}_i - \bar{x}_j)\right]$

then we have the plug- in rule corresponding to the rule that we can write region R_j^* in the Baye's procedure as

R_j^*= {x: h_i (x) ≥h_j (x), j= 1,..., q, j≠i}

= {x: U_{ij}≥ Log (ρ_j/ρ_i), j= 1,…,q, j≠i}

and hence

R_j^*:U_{ij}(x) >log (ρ_j/ρ_i), j= 1,…, q, j≠i

Then we defined that by assign x to π_i if W_{ij} (x)>0 for all j=1,…, k, j≠i for two populations the ML or Z- rule is extended to the present case as follows: Let $d_i(x) = [n_i/(n_i+1)](x-\bar{x}_j)1\sum^{-1}(x-\bar{x}_j)$; i=1,…, k

Then the rule is defined by assign x to π_i if $d_i(x)=\min_j\{d_j(x)\}$ which is also aminimum- distance rule taking account of different sample sizes.

Fisher's Approach:

Suppose that (q) population π_i have the same covariance matrix ∑. suppose that we have the n_i samples from π_i populations. Let us denote the between- group and within- group SS_b matrices by:

$S_b = n_1(\bar{x}^{(1)}-\bar{x})(\bar{x}^{(1)}-\bar{x})+\ldots+n_q(\bar{x}^{(q)}-\bar{x})(\bar{x}^{(q)}-\bar{x})$,

$S_w = (n_1-1)S^{(1)}+\ldots+(n_q-1)S^{(q)}$

Now consider a linear combination of X_i:

$Z= b_1x_1+\ldots+b_px_p = \grave{b}x$

where b= (b_1,…,b_p). The between- group and within- group sums of squares of Z are expressed as $\grave{b}S_bb$ and$\grave{b}S_wb$. As a coefficient vector b with a good separation of q groups, Fisher proposed to maximizing the ratio

$$\ell = \frac{\grave{b}S_bb}{\grave{b}S_wb}$$

Optimum solutions are obtained by considering the characteristic equation

$S_bb_i = \ell_i S_w b_i$, $\grave{b}_i S_w b_i = nS_{ij}$

where$\ell_1 > \cdots > \ell_m \geq 0$are possibly non zero characteristic roots of $S_bS_w^{-1}$ m= min{p, q-1}, δ_{ij} is kronecker'sdalta, and δ_{ii}=1, δ_{ij}=0 for i≠j.

Now we have m (meaningful) functions $Z_i=\grave{b}_ix$, i=1,…,m called canonical discriminant functions.

Know if we extended the decision- theoretic approach for two populations to that for (q) populations, Let $\pi_1,\ldots,\pi_q$ be q populations with density functions $f_1(x),..,f_q(x)$ respectively. we wish to divide the sample space R^p in to q mutually exclusive and exhaustive regions $R_1,\ldots,R_q$.

If an observation falls in to R_i, we say that it comes from π_i. Let c(i/j) be aloss when a π_j observation is assigned to R_i, and c(i/i)=0.

Then the risk is

$$\sum_{i=1}^{q}\sum_{j\neq i} c(j/i)\rho_i p(j/i)$$

and for an equalvar- covariance $\sum_1=\sum_2=\sum_3=\ldots=\sum$ we take three cases $\sum_1=\sum_2=\sum_3=\sum$

$$U(x)=\mathrm{Log}\left(\frac{f(x_1,\theta)}{f(x_2,\theta)}\right)$$

$= B_o + \grave{B}_1 x$

$= (\mu_1-\mu_2)\sum^{-1}[x-\frac{1}{2}(\mu_1-\mu_2)]$

where

$B_o=\frac{-1}{2}(\mu_1-\mu_2)\sum^{-1}(\mu_1+\mu_2)$

$B_1=\sum^{-1}(\mu_1-\mu_2)$

Know the same rule for three variables x_1, x_2, x_3

$u(x)=\mathrm{Log}\,[\left(\frac{f(x_1,\theta)}{f(x_2,\theta)}\right)/f(x_3,\theta)]$

$u(x)=\mathrm{Log}\,f(x_1,\theta)-\mathrm{Log}\,f(x_2,\theta)-\mathrm{Log}\,f(x_3,\theta)$

$u(x)=(\mu_1-\mu_2)\sum^{-1}[x-\frac{1}{2}(\mu_1+\mu_2)]-\mathrm{Log}\,f(x_3,\theta)$

$=(\mu_1-\mu_2)\sum^{-1}[x-\frac{1}{2}(\mu_1+\mu_2)]+\frac{1}{2}\grave{x}\Sigma^{-1}x-\grave{x}\Sigma^{-1}\mu_3+\frac{1}{2}\grave{\mu}_3\Sigma^{-1}\mu_3$

$=(\mu_1-\mu_2-\mu_3)\sum^{-1}[x-\frac{1}{2}(\mu_1+\mu_2+\mu_3)]+\frac{1}{2}\grave{x}\Sigma^{-1}x$

∴The discriminant functions of the three variables

$=\grave{x}\,\Sigma^{-1}(\mu_1-\mu_2-\mu_3)+\frac{1}{2}\grave{x}\Sigma^{-1}x$

know we saw that in the two populations if:

$\frac{f(x_1,\theta)}{f(x_2,\theta)}\geq k\Rightarrow x$falls in the first population

$\frac{f(x_1,\theta)}{f(x_2,\theta)}< k\Rightarrow x$falls in the second population

But here we have two k_9 k_1&k_2

The first onedelate to divide $f(x_1,\theta)/f(x_2,\theta)$ and the second delate to divide $f(x_1,\theta)/f(x_3,\theta)$

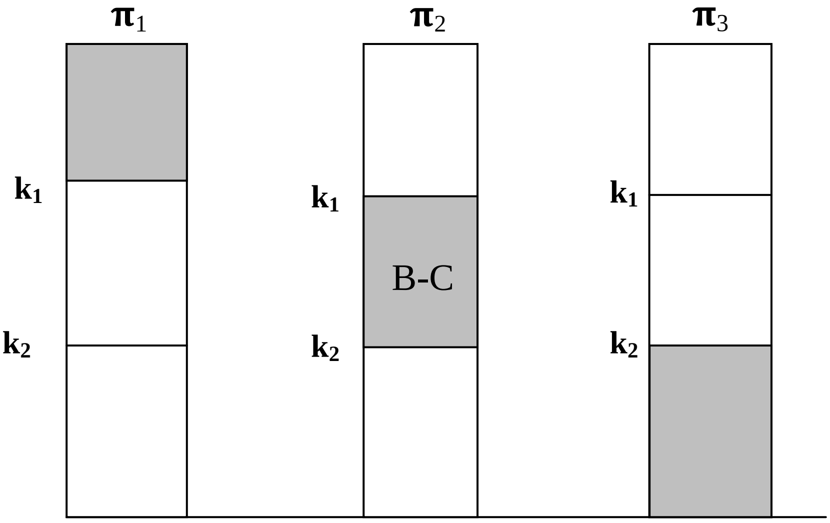
$\frac{f(x_1,\theta)}{f(x_2,\theta)}\geq k_1\Rightarrow x$falls in π_1

$\frac{f(x_1,\theta)}{f(x_2,\theta)}< k_1\Rightarrow x$falls in π_2

$\frac{f(x_1,\theta)}{f(x_3,\theta)}\geq k_2\Rightarrow x$falls in π_1

$\frac{f(x_1,\theta)}{f(x_3,\theta)}< k_2\Rightarrow x$falls in π_3

we have the following:



A:$\frac{f(x_1,\theta)}{f(x_2,\theta)}\geq k_1$

(A∩ C)= x falls in π_3 if $x\geq k_1$

B: $\frac{f(x_1,\theta)}{f(x_2,\theta)}< k_1$ Then x falls in π_2

B – C = x falls in $\pi_2$if $k_2<x<k_1$

C: $\frac{f(x_1,\theta)}{f(x_2,\theta)}\geq k_2$

D: $\frac{f(x_1,\theta)}{f(x_2,\theta)}< k_2$ Then x falls in π_3

Now return to covariance matrix we suppose that:

$$S=\begin{bmatrix} S_{11} & S_{12} & S_{13}\\ S_{21} & S_{22} & S_{23}\\ S_{31} & S_{32} & S_{33}\end{bmatrix}$$

For two groups and for three variables. Then we write the covariance matrix between S and A which is the variance of the third group and denote to it by L, since the variance- covariance between the group A and B and group C is:

Citation: Khawla Mustafa Sadek and BasimaIdris Ismail, 2018. Discriminant Function Using Fisher's Approach For Three Multivariate Normal Populations. Australian Journal of Basic and Applied Sciences., 12(4): 19-21.

$$L=\begin{bmatrix}\ell_{11} & \ell_{12} & \ell_{13}\\ \ell_{21} & \ell_{22} & \ell_{23}\\ \ell_{31} & \ell_{32} & \ell_{33}\end{bmatrix}$$

we use it to calculate "fisher discriminant function"

$$Z=\grave{x}L^{-1}\left(\bar{x}_1^{(i)}-\bar{x}_2^{(i)}-\bar{x}_3^{(i)}\right)+\frac{1}{2}\grave{x}L^{-1}x$$

$$=(x_1\ x_2\ x_3)\begin{bmatrix}\ell_{11} & \ell_{12} & \ell_{13}\\ \ell_{21} & \ell_{22} & \ell_{23}\\ \ell_{31} & \ell_{32} & \ell_{33}\end{bmatrix}^{-1}\begin{bmatrix}\bar{x}_1^{(i)} & -\bar{x}_2^{(i)} & -\bar{x}_3^{(i)}\\ \bar{x}_1^{(i)} & -\bar{x}_2^{(i)} & -\bar{x}_3^{(i)}\\ \bar{x}_1^{(i)} & -\bar{x}_2^{(i)} & -\bar{x}_3^{(i)}\end{bmatrix}$$

$$+\frac{1}{2}(x_1x_2x_3)\begin{bmatrix}\ell_{11} & \ell_{12} & \ell_{13}\\ \ell_{21} & \ell_{22} & \ell_{23}\\ \ell_{31} & \ell_{32} & \ell_{33}\end{bmatrix}^{-1}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

Conclusion:

Fisher's Approach us Baye's method which is reduced to the maximum likelihood classification rule which is use to discriminant about two population can be extend to acted with three or more populations with certain sample space.

This approach is very usefull when we use the quntitivevaribles having acertain chooses such that good, very good, exlent or… ete.

Then we can say that fisher's approach is very usefull in our life to act with many different ways.

REFERENCES

Anderson, T.W., 2003. "Introduction to Multivariate Statistical Analysis" 3rded, Wiley- Hoboken, N. J. (41, 161, 320, 326, 6, 334, 511).

Arnold, S.F., 1981. "The Theory of Linear Models and Multivariate Analysis", Wiley- Hoboken, N. J. (154).

Kendall, M.G., 1966. "Discrimination and Classification, In Multivariate Analysis" (P. R. Krichnaiah, ed.). Academic press, New York (249).

Rao, C.R., 1973. "Linear Statistical Inference and It's Applications", 2nded, Wiley, New York, (312, 503, 505).

Srivatava, M.S., 2007. "Multivariate Theory Analyzing High Dimensional Data", J. Japan statist. Soc: 37, 53-86. (279, 280).

Takeuchi, K. and A. Takemura, 1988. "Some Results on Univariate and Multivariate Cornish- Fisher Expansion: Algehraic Properties and Validity" Sankhyā A, 50, 111-136. (130).

Yasunori Fujikoshi and Vladimir V. Ulyanov and Ryoichi Shimizu, 2011. " Multivariate Statistics High- Dimensional and Large- Sample Approximations".