

# 白鲫与鲫鱼形态鉴别函数\*

蔡庆华

(中国科学院水生生物研究所, 武汉)

**提要** 作为传统分类学方法的补充, 本文以白鲫和鲫鱼的四个外形指标(全长、体长、体高和体重)进行Fisher多元线性判别分析, 建立鉴别此二亚种的形态判别函数, 以便对活体进行直接鉴定。检验表明, 所得结果显著性极高, 鉴别符合率达100%, 且在渔业应用上简单易行。此外, 本文还就在生物学中应用多元判别分析的一般方法作了简明扼要的阐述。

## 一、引言

白鲫(*Carassius auratus curieri* Temminck et Schlegel)又称阪鲫, 原产于日本琵琶湖, 是鲫鱼的一个亚种, 称为源五郎鲫。1897年日本大阪府河内郡一渔民将捕获的源五郎鲫驯养于池中, 经多年选育而成纯系品种并取名为河内鲫, 因其个体大, 生长快, 成为日本各地深受欢迎的主要鲫鱼养殖品种。1959年我国台湾省引进河内鲫, 1973年又引入香港, 经饲养证明是一种高产品种, 现已广泛养殖, 对天然水体鲫鱼资源增殖以及池塘鱼产量的提高都具有一定效果。河内鲫自1976年3月由广东省首次引进我国大陆以来, 因其体色银白, 遂定名为白鲫, 又因其易于繁殖, 病害较少, 起水率高和食性杂等优点, 在我国渔业生产实践中显示出良好的增产效益而具有很高的推广价值。

有关白鲫和鲫鱼的指名亚种普通鲫鱼 *Carassius auratus auratus* (Linnaeus) 的生物学特性、自然条件和人工环境下的食性、生长以及养殖技术方面的研究工作, 已有较系统的报道<sup>[1-4]</sup>。然而, 由于白鲫和普通鲫鱼同属鲤亚科鲫属鱼类, 白鲫除体色较普通鲫鱼稍白之外, 各方面的外观形态特征都与普通鲫鱼很相似。在渔业生产实践中, 特别是在选种育种中要进行活体鉴别, 则有一定困难。因为采用传统分类学方法要准确地鉴别白鲫和鲫鱼时, 需要摘取鳃片计算鳃耙数目, 必然会造成鱼体死亡, 不利于渔业生产。因此, 本文收集一些已经分类学鉴别的白鲫和鲫鱼的数据, 尝试用数学方法寻找一个定量指标(鉴别函数), 以外观可量性状来鉴别白鲫和鲫鱼。

## 二、材料和方法

1985年6月21日和22日两天, 在乌鲁木齐市塔桥湾水库捕得30尾鱼。经分类学鉴定(鳃耙数目37—54者为鲫鱼, 鳃耙数目102—120者为白鲫), 知其中有12尾是白鲫(1龄), 有18尾是鲫鱼(1龄)。对这30尾鱼分别测量全长(cm)、体长(cm)、体高(cm)和体重

\* 本文在导师刘建康教授以及梁彦龄教授的指导下完成。承陈宜瑜教授审阅指导, 刘伏泉、黄尚务先生提供数据和帮助, 特此深表谢意。

(0.5kg)等四项外观形态指标,使用Fisher线性判别分析,对这两组鱼进行处理,建立用以鉴别白鲫和鲫鱼形态判别函数,比较直观地用数量化的外观形态学指标进行定种。

判别分析(discriminant analysis)是判定个体所属类别的一种多元统计方法,它在生物学、生态学上有着广泛的应用,如根据细菌的形态及生化特性判别属于哪一菌种<sup>1)</sup>,根据水体的环境化学指标判断是否出现水华<sup>2)</sup>,以及根据外形特征判别下海鳗鲡的雌雄<sup>[5]</sup>等。判别分析方法是在已知要判别的类型和数目并已取得各种类型的一批已知样本的情况下,根据一个未知样品的多种性质(如化学元素的含量)而判定它究竟属于哪一类。如果判别的类型只有两个(如水华的出现与否)称为二组判别,如果判别的类型有多个(如水体中磷细菌、硝化菌、硫细菌等的判别)叫多组判别。在判别分析中,因判别的准则不同,又有Fisher和Bayes等不同的判别分析法。Bayes判别准则是以个体归属某类的概率(或某类的判别函数值)最大或错分总平均损失最小为标准;而Fisher判别准则是要求各类之间的变异尽可能地大,而各类内部的变异尽可能地小,其中变异性用离均差平方和表示。本文是以Fisher判别准则为依据,建立白鲫与鲫鱼的二组线性判别函数。

Fisher二组判别分析的直观意义就是:如果两类事物的某些指标很相似,则仅凭某一指标往往不能区分这两类事物(如仅用体重或体色就无法分辨出白鲫与鲫鱼),因为两类事物的某一指标观测值常会出现程度的重叠。但若将几个指标综合起来,用数学方法找出这些指标之间的一个适当的线性组合,就可以使重叠程度大为减少,从而比较正确可靠地分辨出这两类事物。这个线性组合可以作为一个综合指标,用以区分两类事物,因而又叫判别函数。即:按Fisher准则作出对多个测量值 $x = (x_1, x_2, \dots, x_m)$ 的函数 $f(x)$ ,并定出一个值 $c$ ,对任一待判别样本的一组观测值 $x_0 = (x_1', x_2', \dots, x_m')$ 代入 $f(x)$ ,若 $f(x_0) \geq c$ ,则判断为属于某一类,而 $f(x_0) < c$ 则判断为另一类。这里 $f(x)$ 称为判别函数, $c$ 称为分界点(或临界判别值)。

假设现有 $n$ 个观测样本,其中A类样本有 $n_1$ 个,B类样本 $n_2$ 个( $n_1 + n_2 = n$ ),每个样本有 $m$ 项观测指标,Fisher二组判别分析的具体步骤是<sup>1)</sup>:

(1) 对观测数据 $x_{ijk}$  ( $k=1, 2, i=1, 2, \dots, n_k, j=1, 2, \dots, m$ ) 求各指标均数

$$\bar{x}_{ik} = \frac{1}{n_k} \sum_{j=1}^{n_k} x_{ijk} \quad k=1, 2, i=1, 2, \dots, n_k$$

(2) 计算类内离均差平方和

$$w_{ij} = \sum_{g=1}^2 \left( \sum_{k=1}^{n_g} x_{kig} \cdot x_{kjg} - n_g \cdot \bar{x}_{ig} \cdot \bar{x}_{jg} \right) \quad i, j=1, 2, \dots, m$$

(3) 解方程组,求判别函数的系数

$$\sum_{j=1}^m w_{ij} \cdot b_j = (n-2) \cdot (\bar{x}_{i1} - \bar{x}_{i2}) \quad i=1, 2, \dots, m$$

1) 李勤生等,东湖异养细菌群落的分类结构和聚类分析,水生生物学报(待发表)。

2) 蔡庆华,武汉东湖浮游植物水华的多元分析,水生生物学报(待发表)。



(4) 建立判别函数  $f(x) = \sum_{i=1}^m b_i \cdot x_i$

确定临界判别值  $c = \frac{z_1 + z_2}{2}$

其中,  $z_1 = \sum_{i=1}^m b_i \cdot \bar{x}_{i1}$        $z_2 = \sum_{i=1}^m b_i \cdot \bar{x}_{i2}$

用统计方法来作判别分析必然会有错误分类的情况发生, 因此需要检验判别函数的判别效果, 其公式为

$$F = \frac{(n_1 + n_2 - m - 1) \cdot n_1 \cdot n_2}{m \cdot (n_1 + n_2 - 2)(n_1 + n_2)} \cdot D^2$$

其中:  $D^2 = \sum_{i=1}^m b_i \cdot (\bar{x}_{i1} - \bar{x}_{i2})$  为Mahalanobis距离

求得F值后, 查表  $F_0 = F_\alpha(m, n-m-1)$ , 若  $F > F_0$ , 则认为变量的判别效果有显著性, 即建立的判别函数有效。

### 三、结果与讨论

上述过程用BASIC语言编制程序, 在IBM—PC/XT微型计算机上运行, 得到两组数据的基本统计量(均值, 方差, 标准差, 最小值和最大值)以及复相关矩阵和协方差矩阵, 同时也得到了总体统计量(均值, 标准差, 样本数)以及协方差矩阵和线性判别函数的各项系数, 由此可知其线性判别函数是:

$$f(x) = 2.47TL - 7.52BL + 7.07BH + 97.69BW$$

其中: TL为全长(cm), BL为体长(cm), BH为体高(cm), BW为体重(0.5kg), 而临界判别值  $c = 15.34$ 。因此, 对任一个待判别的个体, 测其全长、体长、体高和体重之后, 代入  $f(x)$ , 若  $f(x)$  的计算值大于15.34, 则判断为鲫鱼, 否则为白鲫。

检验显著性:  $D^2 = 21.9946$ ,  $F = 35.35^{***}$

可以知道, 此判别函数是有效的。

继而, 还可以检验误判率或判别符合率。将已知的30尾鱼的4项实测指标数据回代到  $f(x)$  之中, 结果发现: 第一组(白鲫)误判为第二组(鲫鱼)的误判率为0, 而第二组误判为第一组的误判率也为0。因此, 本文指出的判别函数对这30尾白鲫与鲫鱼的判断准确率为100%, 是相当精确可靠的。如果用这30尾鱼之外的别的几次采集数据, 如用6月27日在紫湖水库和7月2日在大海子水库捕获的10尾鲫鱼(1龄5尾, 2龄3尾, 4龄2尾, 经分类学鉴定)以及7月27日在塔桥湾水库捕获的5尾白鲫(1龄, 经分类学鉴定)共15尾的上述4项实测指标数据, 代入判别函数  $f(x)$  进行检验, 发现白鲫全部判别正确, 而鲫鱼只有

1) 蔡庆华, 武汉东湖浮游植物水华的多元分析, 水生生物学报(待发表)。

两例误判的情况发生。究其原因,发现误判的那两尾鲫鱼都是4龄的,而其他8尾1~2龄的鲫鱼判别正确。由于用于建立判别函数的白鲫和鲫鱼都是1龄的,因此,本判别函数对4龄鱼无效是可以理解的,换言之,本函数方程的适用范围是1龄左右。这也是在今后的应用中要注意的地方。此外,如有足够的材料,选用同样的方法,也可以建立起适合于鉴别白鲫和鲫鱼的各年龄组的函数方程。

### 参 考 文 献

- [1] 丁端华,池养条件下银鲫与鲫鱼的生物学特性比较及其在生产上的意义,水生生物学集刊,6(2),163—176,1977。
- [2] 丁端华,池养白鲫的食性和生长及其养殖的研究,水产学报,12(4),291—302,1988。
- [3] 中山大学生物系,介绍一种优良鲫鱼品种——白鲫的生物学和养殖,淡水渔业,(7—8),15—21,1977。
- [4] 中山大学生物系,白鲫鱼种对水绵消化吸收率的研究,水生生物学集刊,6(4),409—418,1979。
- [5] 林鼎等,鳊 鲮繁殖生物学研究: I,下海鳊鲮雌雄性状差异和鉴别,水生生物学集刊,6(2),177—190,1977。
- [6] 王学仁,地质数据的多变量统计分析,科学出版社,1982。
- [7] 杨树勤主编,中国医学百科全书——医学统计学,上海科学技术出版社,1985。
- [8] 黄正南,医用多因素分析及计算机程序(第二版),湖南科学技术出版社,1986。
- [9] 黄秉宪等译,计量医学,上海科学技术出版社,1984。

## THE MORPHOLOGICAL DISCRIMINANT FUNCTION FOR THE IDENTIFICATION OF *CARASSIUS* *AURATUS CURIERI* TEMMINEK ET SCHLEGEL AND *CARASSIUS AURATUS AURATUS* (LINNAEUS)

Cai Qinghua

(Institute of Hydrobiology, Academia Sinica, Wuhan)

### Abstract

By using multivariate discriminant analysis of 4 morphological characteristics (total-length, body-length, body-high, body-weight), the morphological discriminant function for the identification of *Carassius auratus curieri* T. et S. and *Carassius auratus auratus* (L.) is yielded. Instead of general taxonomic method, this mathematical method is simple and easy to be applied in fishery practise. The function has been tested as of high significancy, with 100% correct rate. A brief introduction to the use of multivariate discriminant analysis in biology is also presented.