

# 多元分析方法分辨细齿 小塔塔鼠的下臼齿

黄学诗 王令红

(中国科学院古脊椎动物与古人类研究所)

**关键词** 内蒙古 乌兰塔塔尔 啮齿类 多元分析

## 内 容 提 要

啮齿类梳趾鼠科小塔塔鼠的下臼齿的形态非常接近。本文根据已有的一百多个牙齿,用齿冠的长度和宽度两个性状,在坐标系中画出了  $M_1$ 、 $M_2$  和  $M_3$  的一系列椭圆,代表着这些牙齿一定比例的样本分布范围,并以此得出了细齿小塔塔鼠这三种牙齿的判别线。

步林 (Bohlin) 在 1946 年就曾指出,塔塔鼠 (*Tataromys*) 的上下臼齿在大多数情况下,或至少经常是很相似的,所以人们在处理单个牙齿时必须很小心,不要混淆它们。本文前一作者在研究内蒙阿左旗乌兰塔塔尔的梳趾科啮齿类化石过程中,发现塔塔鼠和小塔塔鼠等属的上、下臼齿之间的齿冠形态有一定的差别,尚易区分。但上、下颌骨上的连续三个臼齿,它们彼此之间的尖脊构造却非常接近,依照形态很难区分它们。如果这些牙齿不连接在颌骨上,而单个出现,我们有时只能用前后接触面作些判断。但在某些情况下接触面也不很清楚,这就大大地影响了判别的准确性。如果牙齿所在位置不能正确确定,很自然就影响了标本的适当归属。但我们在观察过程中,发现上下左右每半边上的三个牙齿,在长宽大小上确实存在一定的差异。如在同一齿列上,后面的牙齿一般比前面的较长、较宽。这使我们想到能否用多元数理统计方法去区分它们。本文仅想就此做一尝试。

## 一、材料和方法

材料是 1978 年我所内蒙野外队在内蒙阿左旗乌兰塔塔尔地区采集的细齿小塔塔鼠 (*Leptotataromys gracilidens*) 的下颌骨数十件,其中带有  $M_1$  30 个,  $M_2$  55 个,  $M_3$  52 个。这些牙齿全部连接在下颌骨上,无一脱落,因此它们的位置是正确无疑的。

根据齿冠的长度和宽度分辨下臼齿的原理是从二维随机向量  $(x_1, x_2)$  在平面上的分布确定样本属于某一臼齿 ( $M_1$  与  $M_2$ , 或  $M_2$  与  $M_3$ ) 的概率。其步骤如下:

### 1. 坐标变换

把每类牙齿的长度和宽度的坐标分别旋转一特定的角度,建立新的坐标系统,使每类牙齿的所有样本在新的坐标系统中两个新变量的相关系数等于零。

旋转角度  $\theta$  从下式求出:

$$\operatorname{tg} \theta = \frac{2s_{12}}{s_{11} - s_{22}}$$

$$\text{其中离差阵 } S = (s_{ij}) \begin{cases} s_{ij} = \sum_{a=1}^n (x_{aj} - \bar{x}_j)(x_{ai} - \bar{x}_i) \\ s_{ii} = \sum_{a=1}^n (x_{ai} - \bar{x}_i)^2 \end{cases}$$

$n$  代表例数  $i, j = 1, 2, i \neq j$

## 2. 计算每一样本在新的坐标系 $(Z_1, Z_2)$ 中的数值

每个样本的  $z_1, z_2$  数值从下式中求出:

$$\begin{cases} z_1 = x_1 \cos \theta + x_2 \sin \theta \\ z_2 = -x_1 \sin \theta + x_2 \cos \theta \end{cases}$$

并计算新变量  $Z_1$  和  $Z_2$  的平均数和标准差。

$$\text{平均数 } \bar{z} = \frac{\sum_{a=1}^n z_a}{n}$$

$$\text{标准差 } \sigma_z = \sqrt{\frac{\sum_{a=1}^n (z_a - \bar{z})^2}{n}}$$

$n$  为样本数

## 3. 确定判别线

以  $z_1$  和  $z_2$  的平均数坐标点(也就是原坐标中的平均数坐标点)为中心,以二坐标轴上相同数目的标准差分别作为长、短轴画出若干椭圆,每一椭圆上的点即为二维随机向量的等密度轨迹。而不同种类(在我们的问题中为  $M_1, M_2$  或  $M_3$ )的相同数目的标准差的长、短轴构成的椭圆区域内具有相同数目的样本分布,连接两个种类的这样两个椭圆交点的直线即代表了它们的判别线。

## 二、结 果

### 1. 原始数据

$M_1, M_2$  和  $M_3$  的每一样本的长度 ( $x_1$ ) 和宽度 ( $x_2$ ) 的测量值列于表 1。

### 2. 坐标变换

$$M_1 \text{ 组: } \theta_1 = -43^\circ 12'$$

$$M_2 \text{ 组: } \theta_2 = 44^\circ 52'$$

$$M_3 \text{ 组: } \theta_3 = 26^\circ 18'$$

### 3. 各组样本在经变换后的坐标系统中新变量 $z_1$ 和 $z_2$ 的数值

列于表 2。

### 4. $M_1$ 与 $M_2$ 的分辨

以  $M_1$  的新变量  $Z_1$  和  $Z_2$  的平均数坐标点为中心,以  $4\sigma_{z_1}$  和  $4\sigma_{z_2}$  分别为长、短轴画椭圆。

在  $M_2$  组经变换后的坐标上画出同样的椭圆。

两个椭圆的交点  $A$  和  $B$  的连线即为  $M_1$  与  $M_2$  的判别线(图 1)。

### 5. $M_2$ 与 $M_3$ 的判别

表 1 牙齿测量 (单位: 毫米)

标本号	M <sub>1</sub>		M <sub>2</sub>		M <sub>3</sub>	
	长	宽	长	宽	长	宽
v. 7345. 35	2.7	2.2	3.2	2.8	3.6	2.9
36	2.8	1.7	3.2	2.6	3.3	2.6
37	2.8	1.9	3.1	2.4	3.3	2.2
38	2.7	2.1	3.3	2.7	3.6	2.7
39	2.7	2.0	3.2	2.6	3.4	2.6
40	2.8	2.0	2.9	2.5	3.5	2.5
41	2.5	2.0	3.1	2.2	3.2	2.5
42	2.6	1.9	3.3	2.5	3.3	2.5
43	2.4	1.7	2.8	2.0	2.9	2.0
44	2.8	2.0	3.2	2.3		
45	2.7	2.0	2.9	2.4		
46	2.6	2.1	3.0	2.8		
47			3.0	2.3		
48	2.9	2.4				
51	2.6	2.0	2.9	2.3	2.9	2.2
52	2.7	2.1	3.2	2.4	3.2	2.3
53	2.8	1.9	3.2	2.5	3.2	2.5
54	2.7	2.0	2.9	2.4	2.9	2.4
55	2.9	1.9	3.1	2.4	3.1	2.5
56			3.2	2.4	3.2	2.3
57	2.5	1.5	3.0	2.2	3.2	2.3
58	2.7	2.1	2.7	2.3	3.1	2.4
60	2.7	2.0	3.2	2.4	3.0	2.2
61	2.9	2.2	3.2	2.6	3.1	2.5
62	2.6	2.0	3.0	2.4	2.8	2.4
63	2.7	1.9	3.2	2.4	2.9	2.2
64	2.5	2.0	3.1	2.5	3.1	2.4
65			3.1	2.3	3.4	2.4
66			3.0	2.5	3.3	2.5
67			3.2	2.5	3.4	2.6
68			3.5	2.5	3.5	2.4
69			3.1	2.3	3.2	2.2
70			3.3	2.5	3.3	2.5
71			3.1	2.5	3.2	2.5
72			3.3	2.6	3.4	2.5
73			3.1	2.4	3.1	2.4
74			2.8	2.2	3.1	2.5
75			2.9	2.0	2.9	2.1
76			3.4	2.1	3.4	2.4
77			3.1	2.3	3.3	2.3
78					3.5	2.5
79			3.2	2.5	3.3	2.4
80					3.4	2.3
81			2.9	2.2	3.3	2.6
82			2.8	2.2	3.0	2.2

续表

标本号	M <sub>1</sub>		M <sub>2</sub>		M <sub>3</sub>	
	长	宽	长	宽	长	宽
V·7345·83			3.2	2.4	3.3	2.2
84			3.2	2.3	3.2	2.3
85			3.4	2.5	3.6	2.5
86	2.9	2.0	3.3	2.6		
87			3.2	2.4		
88			3.2	2.5		
89			3.2	2.3		
90			3.1	2.4		
91	3.2	2.1	3.1	2.6		
92	2.9	2.1				
93			3.0	2.5		
94			3.1	2.3		
95					3.4	2.4
96					3.5	2.4
97					3.8	2.6
98					3.9	2.7
99					3.6	2.5
100					3.8	2.4
101					3.6	2.6
102					3.7	2.6
103	2.6	1.8	2.8	2.3	2.6	2.2
104	2.8	2.0				
105			3.4	2.2		
平均数	2.72	1.99	3.11	2.40	3.29	2.42
标准差	0.152	0.1628	0.53	0.1673	0.2678	0.1676

表 2 样本的  $z_1$  和  $z_2$  值

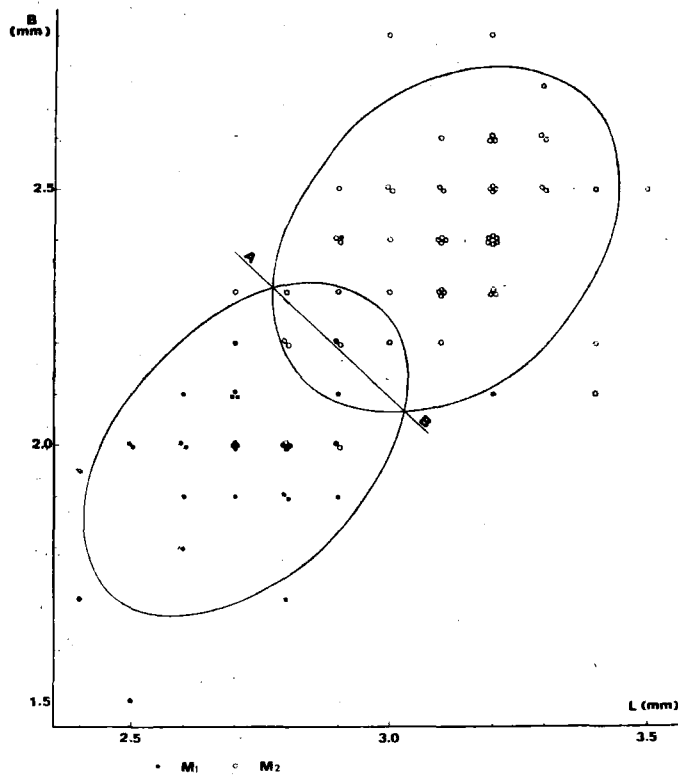
标本号	M <sub>1</sub>		M <sub>2</sub>		M <sub>3</sub>	
	$z_1$	$z_2$	$z_1$	$z_2$	$z_1$	$z_2$
V·7345·35	0.45	3.47	4.26	-0.28	4.52	1.03
36	0.87	3.32	4.12	-0.42	4.11	0.89
37	0.73	3.32	3.90	-0.50	3.93	0.53
38	0.52	3.39	4.26	-0.42	4.43	0.85
39	0.59	3.32	4.12	-0.42	4.34	0.84
40	0.66	3.39	3.84	-0.28	4.25	0.71
41	0.45	3.19	3.76	-0.64	3.98	0.84
42	0.59	3.18	4.12	-0.56	4.07	0.80
43	0.58	2.90	3.41	-0.57	3.49	0.52
44	0.66	3.39	3.90	-0.64		
45	0.59	3.32	3.76	-0.36		
46	0.45	3.32	4.12	-0.14		
47			3.76	-0.50		
48	0.46	3.75				

续表

标本号	M <sub>1</sub>		M <sub>2</sub>		M <sub>2</sub>	
	$z_1$	$z_2$	$z_1$	$z_2$	$z_1$	$z_2$
V. 7345 51	0.52	3.25	3.69	-0.43	3.57	0.70
52	0.52	3.39	3.97	-0.57	3.89	0.66
53	0.73	3.32	4.05	-0.49	3.98	0.84
54	0.59	3.32	3.76	-0.36	3.67	0.88
55	0.81	3.39	3.90	-0.50	3.89	0.89
56			3.97	-0.57	3.89	0.66
57	0.79	2.83	3.69	-0.57	3.89	0.66
58	0.52	3.39	3.55	-0.29	3.85	0.80
60	0.59	3.32	3.97	-0.57	3.66	0.66
61	0.60	3.61	4.12	-0.42	3.89	0.89
62	0.52	3.25	3.83	-0.43	3.58	0.93
63	0.66	3.25	3.97	-0.57	3.57	0.70
64	0.45	3.19	3.98	-0.42	3.85	0.80
65			3.83	-0.57	4.12	0.66
66			3.91	-0.35	4.07	0.80
67			4.05	-0.49	4.20	0.84
68			4.27	-0.71	4.21	0.62
69			3.83	-0.57	3.84	0.57
70			4.12	-0.56	4.07	0.80
71			3.98	-0.42	3.98	0.84
72			4.19	-0.49	4.16	0.75
73			3.90	-0.50	3.85	0.80
74			3.55	-0.43	3.89	0.89
75			3.48	-0.64	3.53	0.61
76			3.90	-0.92	4.12	0.66
77			3.83	-0.57	3.98	0.62
78					4.25	0.71
79			4.05	-0.49	4.03	0.71
80					4.07	0.57
81			3.62	-0.50	4.11	0.89
82			3.55	-0.50	3.66	0.66
83			3.97	-0.57	3.93	0.53
84			3.90	-0.64	3.89	0.66
85			4.19	-0.63	4.34	0.67
86	0.74	3.46	4.12	-0.42		
87			3.97	-0.57		
88			4.05	-0.49		
89			3.90	-0.64		
90			3.90	-0.50		
91	0.89	3.74	4.05	-0.35		
92	0.67	3.53				
93			3.91	-0.35		
94			3.83	-0.57		
95					4.12	0.66
96					4.21	0.62

续表

标本号	M <sub>1</sub>		M <sub>2</sub>		M <sub>3</sub>	
	$z_1$	$z_2$	$z_1$	$z_2$	$z_1$	$z_2$
V. 7345. 97					4.56	0.67
98					4.70	0.71
99					4.34	0.67
100					4.48	0.49
101					4.38	0.76
102					4.47	0.71
103	0.66	3.10	3.62	-0.36	3.30	0.84
104	0.66	3.39				
105			3.97	-0.85		
平均数	0.617	3.333	3.913	-0.498	4.022	0.732
标准差	0.122	0.191	0.199	-0.142	0.297	0.119

图1 M<sub>1</sub> 与 M<sub>2</sub> 的分辨Fig. 1 Chart of discrimination for M<sub>1</sub> and M<sub>2</sub>

以 M<sub>2</sub> 的新变量  $Z_1$  和  $Z_2$  的平均数的坐标点为中心, 以  $2\sigma_{z_1}$  和  $2\sigma_{z_2}$  分别为长、短轴画椭圆。

在 M<sub>3</sub> 组经变换后的坐标上画出同样的椭圆。

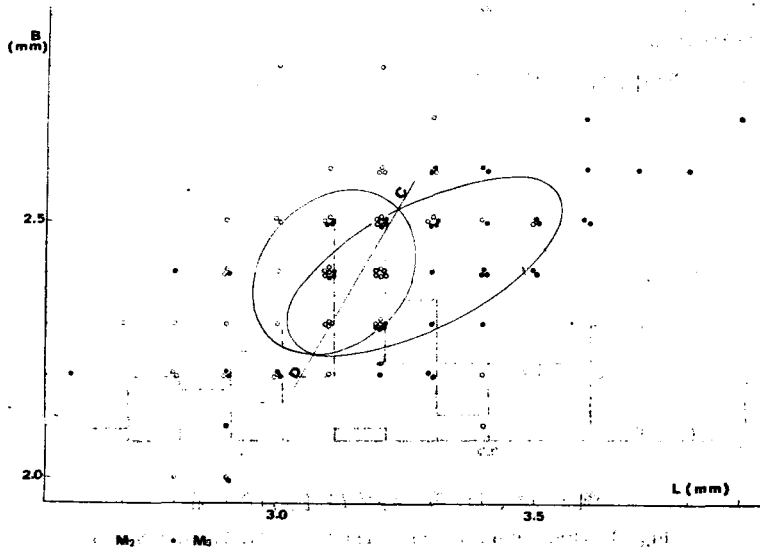


图2  $M_2$  与  $M_3$  的分辨

Fig. 2 Chart of discrimination for  $M_2$  and  $M_3$

这样,两个椭圆的交点  $C$  和  $D$  的连线即为  $M_2$  与  $M_3$  的判别线(图2)。

**6. 回报**

把  $M_1$  和  $M_2$  组每一样本用图1中的判别线  $AB$  重新加以分辨,看其正确与否,分辨效率见表3。

表3  $M_1$  与  $M_2$  的分辨效率

$M_1$			$M_2$			分辨能力
样本数	误判数	正判数	样本数	误判数	正判数	
30	3	27	55	5	50	90.6%

用同样的方法把  $M_2$  和  $M_3$  组中每一样本以图2中的判别线  $CD$  重新加以分辨,分辨效率见表4。

表4  $M_2$  与  $M_3$  的分辨效率

$M_2$			$M_3$			分辨能力
样本数	误判数	正判数	样本数	误判数	正判数	
55	18	37	52	18	34	66.4%

**三、比 较**

为了判断本文使用的二元统计分析方法分辨三种牙齿的效果,现用单个因子的方法

进行同样工作,并将两种方法的结果加以比较。

### 1. 单因子分辨原则

$M_1$ 、 $M_2$  和  $M_3$  三种牙齿在长度上的分布见图 3; 在宽度上的分布见图 4。

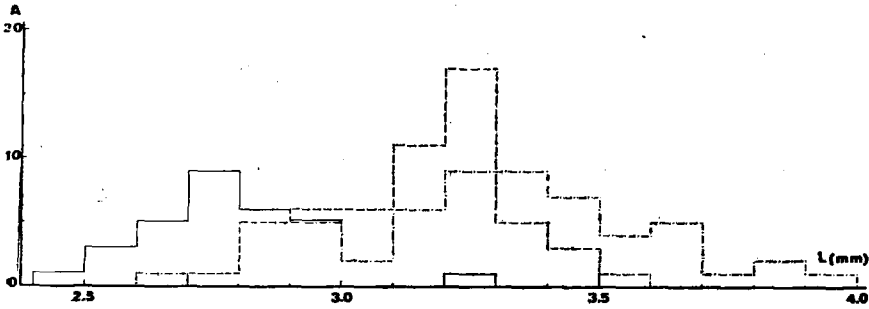


图 3  $M_1$ 、 $M_2$  和  $M_3$  的齿冠长度的分布

Fig. 3 Distribution of crown length of  $M_1$ ,  $M_2$  and  $M_3$ .

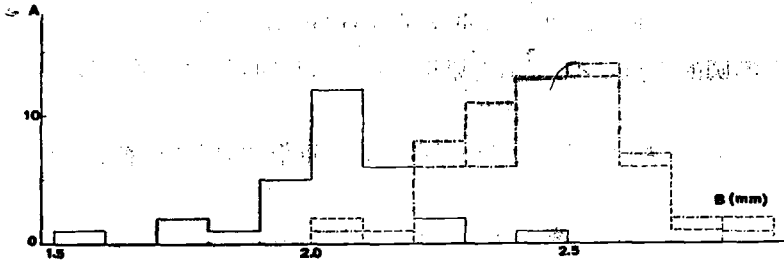


图 4  $M_1$ 、 $M_2$  和  $M_3$  的齿冠宽度的分布

Fig. 4 Distribution of crown breadth of  $M_1$ ,  $M_2$  and  $M_3$ .

以用齿冠长度分辨  $M_1$  与  $M_2$  为例:

判别点在两个平均数之间,并且与它们的距离为相同数目 ( $a$ ) 的各自标准差。

$$a = \frac{|\bar{x}_{M_1} - \bar{x}_{M_2}|}{\sigma_{M_1} + \sigma_{M_2}}$$

用同样的方法,可找出  $M_1$  与  $M_2$  在宽度上的判别点、 $M_2$  与  $M_3$  在长度上和宽度上的判别点。

### 2. 单因子分辨效率

现在,我们用与二元统计分析的同样方法,用每一个样本的测量值进行回报,以确定单因子的分辨效率。单因子误判的情况有两种:一种为两个单因子相同的误判;另一种为两个单因子不同的误判,即是二因子分辨的结果互相矛盾。单因子的分辨效率见表 5。

### 3. 两种分辨能力的对比

在  $M_1$  与  $M_2$  的分辨中,单因子的分辨能力为 83.5%,而二因子的分辨能力为 90.6%。后一方法的效率有所提高。需要提一下的是用单因子分辨时,在 85 例数中有 11 例出现矛盾的判别。也就是说,根据齿冠长度应判为  $M_1$ ,而根据齿冠宽度则应判为  $M_2$ ,或相反。



表 5 单因子的分辨效率

M <sub>1</sub> 与 M <sub>2</sub>			M <sub>2</sub> 与 M <sub>3</sub>		
长度	样本数	85	长度	样本数	107
	正判数	74		正判数	76
	误判数	11		误判数	31
宽度	样本数	85	宽度	样本数	107
	正判数	79		正判数	57
	误判数	6		误判数	50
矛盾判别数		11	矛盾判别数		35
总误判数		14	总误判数		58
分辨效率		83.5%	分辨效率		45.8%

而多元分析的分辨就避免了这类问题的出现。

在 M<sub>2</sub> 与 M<sub>3</sub> 的分辨中, 单因子的分辨能力为 45.8%, 比 50% 还低, 而若分辨率为 50% 的话, 表示根本起不到分辨作用。因此, 可以有把握地说, 用单因子的方法无力分辨这种动物的某个牙齿是 M<sub>2</sub> 还是 M<sub>3</sub>。二元统计的分辨率为 66.4%, 可对待判的 M<sub>2</sub> 或 M<sub>3</sub> 作粗略的估计。但与 M<sub>1</sub> 与 M<sub>2</sub> 的分辨比较起来, 毕竟还是不够理想的。

#### 四、结论与意义

分析结果表明, 用多元统计分析的方法, 即用牙齿的齿冠长度和宽度两个指标同时进行分析, 是可以将细齿小塔塔鼠的三种下臼齿初步地区分开来的(尽管在 M<sub>2</sub> 与 M<sub>3</sub> 的分辨上还不够令人满意)。这比单纯用一个性状(单一的齿冠长度或宽度)进行鉴别要理想得多。

我们根据现有的材料, 一百多个牙齿, 绘出了这个种的三个下臼齿的分布概率图。今后若有该种单个下臼齿的发现, 就可用这个分布图(图 1 或图 2) 大体上确定它的归属。我们只需根据齿冠长度和宽度的数值确定其坐标点位于判别线何侧即可。

国外用多元分析的方法来进行小化石的研究, 已见于多种文献, 并收到很好的效果。但在我国古脊椎动物界, 这项工作才刚刚开始。随着我国脊椎动物学的研究深入, 人们越来越多地注意小哺乳动物的采集和研究, 其中不少类群, 形态差异不大, 只是在个体大小或某些性状上有所不同, 我们也可以用多元分析的方法作为一种手段来划分它们的类群。

此外, 筛洗法也常被用来寻找小化石, 这样, 单个的牙齿就会大量出现。相信, 用多元数理统计的方法来鉴别它们定会发挥越来越大的作用。

(1983 年 6 月 13 日收稿)

#### 参 考 文 献

- 王令红, 1983: 分辨巨猿第一与第二臼齿的多元分析方法。人类学学报, 2(1), 27—39。  
 Agterberg, F. P., 1974: Geomathematics. Elsevier Scientific Publishing Company, Ottawa.  
 Anderson, T. W., 1958: Introduction to multivariate statistical analysis. John Wiley and Sons, New York.  
 Bohlin, B., 1946: The fossil mammals from the Tertiary deposit of Taben-Buluk, Western Kansu. Part 2: Simplicidentata, Carnivora, Perissodactyla and Primates. *Palaeontologia Sinica*, N. S. C., No.

8b, pp. 75—146.

Kendall, M. G., 1975: *Multivariate analysis*. Charles Griffin, London.

## DISTINGUISHING LOWER MOLARS OF *LEPTOTATAROMYS GRACILIDENS* BY MULTIVARIATE ANALYSIS

Huang Xueshi Wang Linghong

(*Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica*)

**Key words** Ulatatal, Nei Monggol; Rodentia; Multivariate analysis

### Abstract

The morphology of the first, second and third lower molars of *Leptotataromys gracilidens* is very similar to one another. It is difficult to distinguish them by any single trait if these preserved as isolated teeth. Over 100 lower molars attached to several dozens of mandibles of this species collected from Ulatatal, Alashan Zuoqi, Nei Monggol are treated in the present paper. Based on the crown length and breadth of these teeth, three series of ellipses in two coordinate charts are drawn, which represent the ranges of different sample distribution probabilities of  $M_1$ ,  $M_2$  and  $M_3$ , respectively. Thus, the discriminative lines between  $M_1$  and  $M_2$  and between  $M_2$  and  $M_3$  are obtained. With these lines, one can recognize preliminarily the position of the isolated lower molars of this species found in the future.