

# 高级生物统计学

ADVANCED BIostatISTICS

马寨璞 编著



科学出版社



2015年河北省生物学强势特色学  
科建设项目资金及2014年中西部  
高校提升综合实力专项资金人才  
培养和创新团队建设项目资助

# 高级生物统计学

马寨璞 编著



科学出版社

北京

## 内 容 简 介

本书主要为《生物统计学》的后续课程内容,对生物科研人员经常遇到的数据分析问题进行了通俗的讲解,并给出了数据分析的通用程序。全书共10章,内容包括多维数据的组织与整理、多元统计学与矩阵计算基础、假设检验、判别分析、聚类分析、主成分分析、因子分析、对应分析、典型相关分析和多维标度法。本书充分考虑了生物科研人员的数学基础,以深入浅出的文字对原理进行了讲解,以极其详细的步骤展示了数据分析过程,以傻瓜式的通用代码给出了解决方案,可以帮助生物科研人员对数据分析做到抓大放小,既领会数据分析的思想,知道为什么要这样做,又方便执行,可以很快得到结果。

本书适合生物类、药学类专业高年级本科生或研究生使用,作为教材或自学材料均可,也是教师和科研工作者的得力助手。

### 图书在版编目(CIP)数据

高级生物统计学/马寨璞编著. —北京:科学出版社,2016.1

ISBN 978-7-03-046802-4

I. ①高… II. ①马… III. ①生物统计 IV. ①Q-332

中国版本图书馆CIP数据核字(2015)第321684号

责任编辑:刘 畅 / 责任校对:贾伟娟  
责任印制:赵 博 / 封面设计:迷底书装

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

文林印刷厂印刷

科学出版社发行 各地新华书店经销

\*

2016年1月第一版 开本:787×1092 1/16

2016年1月第一次印刷 印张:20 1/2

字数:486 000

定价:58.00元

(如有印装质量问题,我社负责调换)

# 前 言

随着实验技术的发展与数据获取的便捷,越来越多的数据出现在科研人员面前,面对庞大的数据,如何从中挖掘出人们所需要的信息,是每一个生物科技工作者感到棘手的问题,虽然有类似《多元统计分析》这样的教材可供参考,但总结起来,这些教材有以下几个特点。

一是多数教材面向统计学专业或者数学专业学生,这类教材有各种译著和影印版等,教材内容多,有些教材更偏重于数学推导,不适合生物学类研究生的“弱”数学背景。

二是多数教材中涉及的例题,其数据都是和财经管理类专业相关,对问题的解释也都是和财经类相关的,对于生物类各专业的学生来讲,这些例题内容显得过于生疏,“不合胃口”。

三是多数教材中使用 SPSS 或者 SAS 软件进行数据分析,讲授 SPSS 操作过程,截图数据解释,有些必要的细节在正文中没有讲授,却在 SPSS 操作中直接解释,显得过于突兀。另外,对于关心数据操作细节,喜欢刨根问底的细节探究者,使用 SPSS 操作,无法显示每一步的具体结果。

为此,在讲授《生物统计学》的基础上,笔者试着以一种新的方式编写这本《高级生物统计学》。

在理论讲授上,做到既有必要的理论知识介绍,又不过分拘泥于数学理论细节,做到理论上理解“是那么回事”即可,不探究数学细节,够用就好。《高级生物统计学》中,每一种方法都有其对应的哲学思考,如主成分分析,我们就认为它对应着哲学中的“抓住主要矛盾”或者“抓住矛盾的主要方面”,在讲授理论的过程中,适当拓展这种方法的哲学应用基础,有助于我们进行创新思考。

但在具体计算实施上,则采用和理论讲授相反的策略,我们利用自编的 MATLAB 代码,具体实现每一步的操作,做到理论讲授的“粗放”和动手计算的“细化”相结合,该粗的要粗,该细的要细。因此对每一个例题,尽量做到详细计算,几乎都给出了每一步的具体计算结果,以方便读者观察、验证。这样安排理论介绍与例题解释,有助于读者更加深刻地掌握知识。

在实践操作与应用上,尽量做到减少输入与步骤。在具体应用时,读者只需要提供格式化的数据,然后参考程序说明写上一两句调用代码即可,每个程序都提供详细操作说明,具有“傻瓜式”、“一键式”操作特点,操作极其简单明了,书中所附每段代码都经过了实践验证,在 MATLAB 2012a 及以上版本运行正常,低于该版本的 MATLAB,有可能个别指令不兼容。

在习题的选择上,每一章的习题都是从众多已经公开发表的文献中精心挑选的,选用这些数据作为对应章节的习题,一是可以让读者进行习题训练,达到掌握本章节基础理论知识的目的;二是拓展读者的思考深度,读者可以研习每道习题的计算结果,并参考原作者是如何解析这些数据的,对比自己的解读,会发现如何才能做到更深入的分析,为此,

每道习题都给出了详细的参考文献出处。本书代码及习题数据可填写本书最后一页的读者反馈表索取。

高级生物统计学是面向生物类各专业研究生的必修专业基础课，其主要内容也是各类生物科研人员必须掌握的数据分析方法。在教学过程中，笔者遇到许多需要更深入学习统计知识的研究生，他们对统计知识的理论要求不高，要求方法操作越简单越好。基于此，本书采用 MATLAB 编程实现每一章的计算，并附于每章的末尾，以帮助学习验证及满足工作需要。另外，对于所附的编程代码，笔者不刻意追求过分精简，毕竟程序是给人读的，过分简练的代码，像 $(^{\backslash}(\backslash\{3\}\backslash)\backslash\{3\}[-]? )? \backslash\{3\}[-]? \backslash\{4\}\$)$ 这样的正则表达式，虽然缩短了程序的长度，节省了篇幅，但类似于天书，并不适合日后的维护。

在本书编写过程中，河北大学生命科学学院院长康现江教授曾多次关心编写进度，并给予鼓励与帮助。对于本书的编写出版，管越强教授也多次给予具体业务指导与帮助。在书稿文字校核时，研究生张凯利和王慧欣做了耐心细致的阅读。科学出版社的编辑对本书的出版付出了辛勤的工作，对于他们的支持与帮助，表示衷心的感谢。本书的出版，还得到了“2015 年河北省生物学强势特色学科建设项目”资金及“2014 年中西部高校提升综合实力专项资金人才培养和创新团队建设项目”资金的支持，在此一并表示深深的感谢。

马寨璞

2015 年 12 月

于河北大学

# 目 录

## 前言

第一章 多维数据的组织与整理	1
第一节 多维数据的组织	1
第二节 多维数据的可视化	2
习题	19
第二章 多元统计学与矩阵计算基础	23
第一节 多维随机变量的一些概念	23
第二节 多元正态分布	25
第三节 WISHART 分布	29
第四节 HOTELLING 分布	30
第五节 WILKS 分布	31
第六节 多元随机变量的数字特征	32
第七节 矩阵分解与导数	33
第八节 数据读取的 MATLAB 实现	35
习题	46
第三章 假设检验	48
第一节 引言	48
第二节 均值向量的检验	49
第三节 协方差矩阵的检验	53
第四节 实例计算	55
第五节 假设检验的 MATLAB 实现	58
习题	72
第四章 判别分析	76
第一节 引言	76
第二节 距离判别法	76
第三节 Bayes 判别法	81
第四节 Fisher 判别法	84
第五节 判别分析的 MATLAB 实现	86
习题	92
第五章 聚类分析	97
第一节 引言	97
第二节 相似性	97
第三节 系统聚类法	101
第四节 系统聚类分析 MATLAB 实现	108

第五节	热图与聚类	117
第六节	其他聚类方法	134
习题		145
<b>第六章</b>	<b>主成分分析</b>	<b>150</b>
第一节	引言	150
第二节	主成分分析的几何解释与推广	151
第三节	数学原理与主要性质	153
第四节	主成分应用举例	156
第五节	主成分分析的 MATLAB 实现	167
习题		176
<b>第七章</b>	<b>因子分析</b>	<b>181</b>
第一节	因子分析的基本概念	181
第二节	因子模型	182
第三节	载荷矩阵的求解	186
第四节	因子旋转与得分	190
第五节	因子分析详细算例	194
第六节	因子分析的 MATLAB 实现	203
习题		209
<b>第八章</b>	<b>对应分析</b>	<b>215</b>
第一节	对应分析的基本思想	215
第二节	对应分析数据的列联表表示	215
第三节	对应分析的基本理论	217
第四节	对应分析的具体步骤与注意事项	220
第五节	对应分析举例与 MATLAB 实现	224
习题		250
<b>第九章</b>	<b>典型相关分析</b>	<b>254</b>
第一节	引言	254
第二节	基本理论性质	254
第三节	样本数据的典型相关分析	260
第四节	典型相关分析的几个问题	263
第五节	典型相关计算过程实例	266
第六节	典型相关分析解读实例	269
第七节	典型相关分析的 MATLAB 实现	276
习题		285
<b>第十章</b>	<b>多维标度法</b>	<b>290</b>
第一节	引言	290
第二节	古典 MDS 的基本原理	291
第三节	非度量 MDS 的求解	295

第四节 权重多维标度.....	297
第五节 古典 MDS 实例计算详解 .....	298
第六节 多维标度的 MATLAB 实现.....	305
习题.....	312
参考文献.....	316



# 第一章 多维数据的组织与整理

## 第一节 多维数据的组织

随着实验技术的提高与实验手段的多样化,生物科研工作者面临越来越多的实验观测数据,如何对这些结果进行有效处理,从而找出新规律、新结论,是每一个生物研究人员必须面对的问题。

在处理多维试验数据时,常常使用多元统计的方法。因此,多维生物试验数据的处理,需要按照多元统计的要求进行整理。表 1-1 是张淑珍等<sup>[1]</sup>研究野生大豆对大豆疫霉根腐病的抗感反应结果,通过观察,可得到常用数据表格的基本特征。

在表 1-1 中,研究数据包含了 5 个指标,分别是来源、叶形、茸毛、粒形、抗感反应,该项试验共进行了 40 次观测。一般的,设  $p$  代表指标(变量)个数,设试验观测次数为  $n$ ,则把这  $p$  个指标表示为  $X_1, X_2, \dots, X_p$ , 常常使用向量

$$X = (X_1, X_2, \dots, X_p)'$$

表示对同一个个体观测的  $p$  个变量。

表 1-1 野生大豆对大豆疫霉根腐病的抗感反应结果

编号	代号	来源	叶形	茸毛	粒形	抗感反应
1	ZYD00225	黑龙江	椭圆	灰色	小形	I
2	ZYD01914	辽宁	椭圆	灰色	小形	I
3	ZYD03803	陕西	椭圆	灰色	小形	I
4	ZYD04437	浙江	椭圆	灰色	小形	I
5	ZYD00434	黑龙江	椭圆	灰色	小形	I
6	ZYD02025	辽宁	椭圆	棕色	大形	I
7	ZYD04320	四川	长叶	灰色	大形	I
8	ZYD03804	陕西	椭圆	灰色	小形	I
9	ZYD02413	辽宁	椭圆	灰色	小形	I
10	ZYD017240	辽宁	椭圆	灰色	小形	I

注: 摘取原文献表中部分数据; “I” 代表抗病

高级生物统计分析和本科期间学习的生物统计学<sup>[2]</sup>一样, 将研究对象的全体称为总体, 表 1-1 中数据有 5 个指标, 称该观测数据的总体为 5 维总体。像表 1-1 这种数据组织格式, 有利于使用数学方法进行处理。一般而言, 当实验数据有  $p$  个指标, 有  $n$  次观测时, 常常排列成如表 1-2 的格式, 表头安排各个指标, 左侧第一列是试验观测次号, 或者是其他诸如样品个体名称之类的属性。

表 1-2 数据的组织形式

次数	$X_1$	$X_2$	...	$X_p$
1	$X_{11}$	$X_{12}$	...	$X_{1p}$
2	$X_{21}$	$X_{22}$	...	$X_{2p}$
...	...	...	...	...
$n$	$X_{n1}$	$X_{n2}$	...	$X_{np}$

取出表 1-2 中的数据, 则  $n$  次观测  $p$  个指标的数据列表形成如下数据矩阵。

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

在数据矩阵的基础上, 可使用其他方法进行深度处理。

## 第二节 多维数据的可视化

在对观测数据进行深入分析处理之前, 有必要先对数据的概况做一些了解。多维数据的可视化, 就是借助数学变换的手段, 以图像的形式表达数据, 更直观地展现出数据具有的特征。

在多维数据的可视化中, 目前有许多可用的方法, 这些方法都是通过将数据进行某种变换, 转化成人们“看起来”更容易理解的形态, 并展示出来。更进一步地讲, 多维数据的可视化, 其实是多维数据“可感知化”的一种, 就是让人更容易感受到、了解更深刻的一种变换手段。这种变换就在身边, 比如持续热播的电视剧《红楼梦》, 其实是对文学著作的可视化, 编剧及制片人等将文字形态的内容, 转化成了可视化的画面。而人类内心的某种感受, 比如苦闷, 也可以通过某种转换, 如通过声音表达成人们容易感受到的状态, 名曲《二泉映月》和《梁祝》等就是这种感受的转换典型。当然, 这种数据的转换超出了生物科研工作者的数据转换范畴, 可以看作更广泛意义上的数据变换。

在高级生物统计学中, 要学习常见的轮廓图法、雷达图法、调和曲线法与星座图方法, 更多的可视化方法, 参阅陈为<sup>[3]</sup>的著作, 切比雪夫脸图参看陆璇<sup>[4]</sup>的译著。为叙述方便, 规定如下: 设  $p$  个指标  $n$  次观测的样本数据矩阵  $\mathbf{X}$ , 则第  $k$  次观测值记为

$$X_{(k)} = (x_{k1}, x_{k2}, \cdots, x_{kp}), \quad k = 1, 2, \cdots, n$$

样本数据矩阵记为

$$\mathbf{X} = (x_{ij})_{n \times p}$$

### 一、轮廓图

轮廓图也可以看作折线图, 它在数据预处理中经常用到, 主要是把每个样本点的各个变量依次用直线连接, 这样就可以看出不同样本点的变化走势, 让人们能够更加清楚地看到差别。这种图形的绘制, 多数以变量作为横坐标, 比如有  $p$  个变量, 则横坐标取  $p$  个点, 而每

一个样本点的一次观测 ( $p$  个变量则有  $p$  个数据), 可分别对应横坐标的  $p$  个点, 其高度可按正比确定量值。每确定一个样本点 (含有  $p$  个横坐标数据), 即可用直线段依次连接各点, 得到一条折线, 即该次观测值的一条轮廓线。对于多个样本点, 每次都重复上述步骤, 可画出多条折线, 构成多个观测值的轮廓图。

这种轮廓线的优点是易于表现变化趋势, 但有时也带来混乱, 比如不相关的各变量之间, 并没有什么变化趋势, 观察这种图, 有时却能够让人“感觉到”某种趋势。轮廓图还常常用来对样本个体进行概要判断归类, 对聚类分析很有帮助。例如, 朱娜<sup>[5]</sup>对渤海湾 7 个不同位置夏季水生动物重金属污染指数进行了测量, 其数据如表 1-3 所示。

表 1-3 夏季水生动物重金属污染指数表<sup>[5]</sup>

采样点	Zn	Cu	Cd	Hg	Pb
S <sub>1</sub>	0.305	0.087	0.117	0.010	0.071
S <sub>2</sub>	0.077	0.044	0.043	0.007	0.138
S <sub>3</sub>	0.132	0.023	0.033	0.007	0.071
S <sub>4</sub>	0.216	0.029	0.046	0.003	0.111
S <sub>5</sub>	0.135	0.010	0.015	0.003	0.015
S <sub>6</sub>	0.141	0.023	0.032	0.003	0.077
S <sub>7</sub>	0.106	0.016	0.028	0.003	0.055

则其轮廓图可绘制为图 1-1。

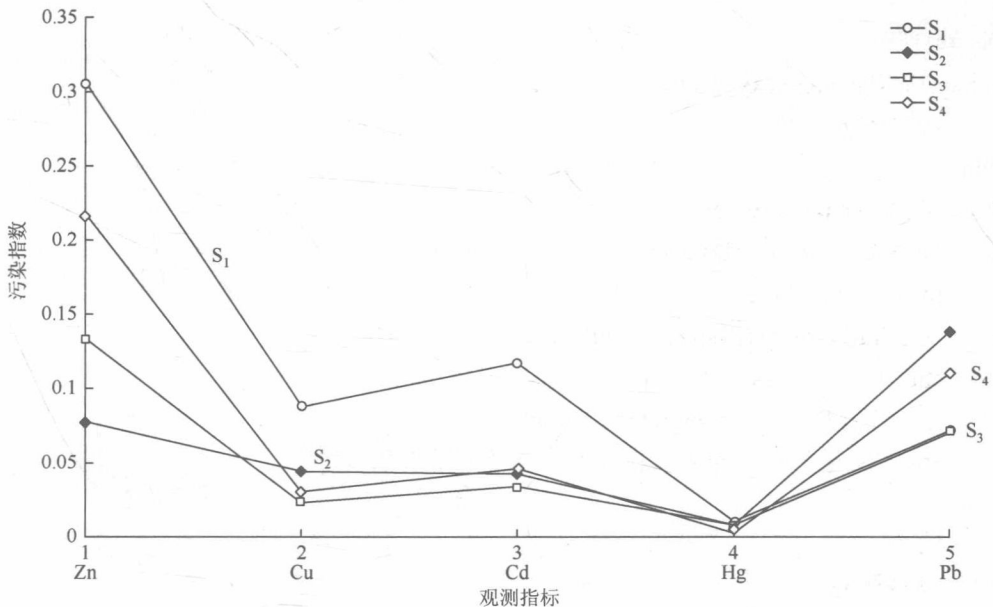


图 1-1 渤海湾不同位置夏季水生动物重金属污染指数 (只给出 4 个点位)

这种轮廓图将两种含量的对比显示得很清楚。下面为笔者编写的 MATLAB 绘图通用代码, 其在 MATLAB 2014a 版本运行正常。

```

function DrawSketch(ExpData,xyLabTxt,VarName)
%用途：对于给定的数据，绘制轮廓图
%参数：本函数需要3个输入参数，ExpData,xyLabTxt,VarName，分别保存数据、
%      坐标轴名称和变量名称。其中数据参数ExpData不可缺省。
% ExpData: 数据以矩阵的形式读入，每行作为一个样本点，每列作为一个变量指标。
%      具体格式如下：
%          0.305,0.087,0.117,0.010,0.071;
%          0.077,0.044,0.043,0.007,0.138;
%          0.132,0.023,0.033,0.007,0.071;
% xyLabTxt: 用来标注绘图坐标轴的文字，属于固定文字，具体格式如下：
%          xyLabTxt={'X 坐标轴','Y 坐标轴'};
% VarName: 用来标注横坐标轴上的变量的名称，属于固定文字，格式如下：
%          VarName={'含量','浓度','渗透率','可视面积'};
%编制：2011.01.17.
%联系：mazhaipu, wdwsjlxx@163.com
%

if nargin>3
    error('Too many Inputs!');
end

% 处理坐标轴标注
if nargin<2||isempty(xyLabTxt)
    xyLabTxt={'X','Y'};
end
if nargin<3||isempty(VarName)
    VarNum=size(ExpData,2);
    for ic=1:VarNum
        VarName{ic}=sprintf('Var%d',ic);
    end
elseif size(ExpData,2)~=length(VarName)
    error('变量名称个数与数据变量个数不符合!');
end

%% 图像配置
FigX0 = 0.05;          % 整图坐标起点 x0
FigY0 = 0.05;          % 整图坐标起点 y0
FigWidth = 0.80;      % 整图图幅宽度
FigHeight = 0.80;     % 整图图幅高度

```

```

% 绘图区位置
CentralAxesX0 = 0.2;           % 坐标起点
CentralAxesY0 = 0.2;           % 坐标起点
CentralAxesWidth = 0.6;        % 图幅宽度
CentralAxesHeight = 0.6;       % 图幅高度
% 将图面分成 2 部分
FigPos = [FigX0, FigY0, FigWidth, FigHeight];
H=figure('Color',[1,1,1], 'Units','normalized',...
         'Position', FigPos, 'Name','HeatMap',...
         'NumberTitle','on');
axis off;

[nRows, nCols]=size(ExpData);
if nRows<=3
    ColorArray=[1,0,0;0,1,0;0,0,1];
else
    ColorArray=rand(nRows,3);
end
% 标注
LegTxt=cell(1,nRows);
MarkNum=13;           % 固定使用 13 种不同标记
MyMarker={'o','*','s','d','v','>','<','p','h','x','!','^','+'};
LineStyleNum=4; % 固定使用 4 种线型
MyLineTypes={'-',':','--','-.'};
myFontSz=18;

for ir=1:nRows
    % 不同标记的循环使用
    MarkPos=mod(ir,MarkNum);
    if MarkPos==0
        MarkPos=MarkNum;
    end
    % 不同线型的循环使用
    LinePos=mod(ir,LineStyleNum);
    if LinePos==0
        LinePos=LineStyleNum;
    end
    plot(1:nCols,ExpData(ir,:),'Color',ColorArray(ir,:),...
        'Marker',MyMarker{MarkPos}, 'LineWidth',1.5,...

```

```

        'LineStyle',MyLineTypes{LinePos});
    Pos4X=ir*(nCols-1)/nRows+1;           % 横坐标位置的确定
    LeftPos4X=floor(Pos4X);                % 为插值计算 y, x1 的位置
    RightPos4X=LeftPos4X+1;              % 为插值计算 y, x2 的位置
    if RightPos4X>nCols;
        RightPos4X=nCols;
    end
    ratio=Pos4X-floor(Pos4X);              % 插值比例
    LegTxt{ir}=sprintf('S%d',ir);        % 图例文字
    Pos4Y=triMix(ExpData(ir,LeftPos4X),ExpData(ir,RightPos4X),ratio);
    text(Pos4X,Pos4Y,LegTxt(ir),'Color','k',...
        'FontSize',myFontSz,'horizontalAlignment','left');
    hold on;
end

set(gca,'xTick',[1:nCols],'FontSize',myFontSz,...
    'FontName','Times',...
    'LineWidth',1.5);
% XY 轴标注
tempX=get(gca,'xtick');
xTickDelta=tempX(2)-tempX(1);
xpx=nCols+0.1*xTickDelta;
tempY=get(gca,'ytick');
ypx=min(tempY,[],2);
xlabel(xyLabTxt{1},'FontName','Times',...
    'rotation',0, 'HorizontalAlignment','Left',...
    'Position',[xpx,ypx], 'FontSize', myFontSz);
xpy=min(tempX,[],2);
yTickDelta=tempY(2)-tempY(1);
ypy=max(tempY,[],2)+0.1*yTickDelta;
ylabel(xyLabTxt{2},'FontName','Times',...
    'rotation',0, 'HorizontalAlignment','Right',...
    'Position',[xpy,ypy], 'FontSize', myFontSz);
legend(LegTxt,'Location','BestOutside');
box off;

%% 底部区域
BottomPos =[ FigX0, FigY0, FigWidth, CentralAxesY0-FigY0];
HD=axes(...

```

```

'Position',BottomPos,...
'Color',[1,1,1],'Box','off',...
'Units','normalized',...
'FontName','times ',...
'FontAngle','italic',...
'FontSize',myFontSz);
axis off;
txtArea=get(HD,'Position'); %[x,y,w,h]
detX4b=txtArea(3)/(nCols-1);
for ic=1:nCols
    xPos4Var=txtArea(1)+(ic-1)*detX4b;
    yPos4Var=txtArea(2);
    text(xPos4Var,yPos4Var,VarName(ic),'Color','k',...
        'rotation',70,...
        'FontSize',myFontSz);
    hold on;
end

%% 子函数
function [mx]=triMix(x1,x2,ratio)
% 计算线性插值,返回 mx
% ratio--介于 x1 和 x2 之间的比例系数,以距离 x1 为标准,该值不大于 1。
delta=x1-x2;
if delta>0
    mx=(1-ratio)*delta+x2;
elseif delta<0
    mx=x1+ratio*abs(delta);
else
    mx=x1;
end

```

## 二、雷达图

雷达图应用很广,例如,在临床眩晕、平衡障碍的诊断中,常用重心动摇检查计检查机体的平衡功能。测量的指标(变量)分别为外周面积、单位面积轨迹长、 $x$ 轴动摇平均中心偏移、 $y$ 轴动摇平均中心偏移及罗姆伯格率。每一指标对诊断具有不同的意义,表 1-4 给出了重心动摇检查计测量的 5 个指标,根据这 5 个指标可制作 5 轴雷达图。这种图形既像雷达荧光屏上看到的图像,也像一个蜘蛛网,因此有人称为雷达图,也有人称为蜘蛛图。

表 1-4 重心动摇检查计测量的 5 个指标<sup>[6]</sup>

指标	均数	标准差	例 1	例 2	例 3
外周面积/cm <sup>2</sup>	3.61	1.86	2.56	36.28	26.09
单位面积轨迹长/(cm/cm <sup>2</sup> )	27.70	9.81	28.76	11.56	28.88
x 轴动摇平均中心偏移/cm	0.07	0.58	0.07	1.63	1.08
y 轴动摇平均中心偏移/cm	-0.32	1.83	-1.44	-0.17	4.07
罗姆伯格率/(cm <sup>2</sup> /cm <sup>2</sup> )	1.39	0.53	1.09	3.99	1.18

一般的, 雷达图(图 1-2)的作图步骤<sup>[7]</sup>如下: ①作圆, 将圆周分为  $p$  等分; ②连接圆心和各分点, 以  $p$  条半径作为各变量坐标轴, 并标注合适刻度; ③将各样本点的  $p$  个变量测定值标记在相应坐标轴, 每个样本点数据连接成一个  $p$  边形。

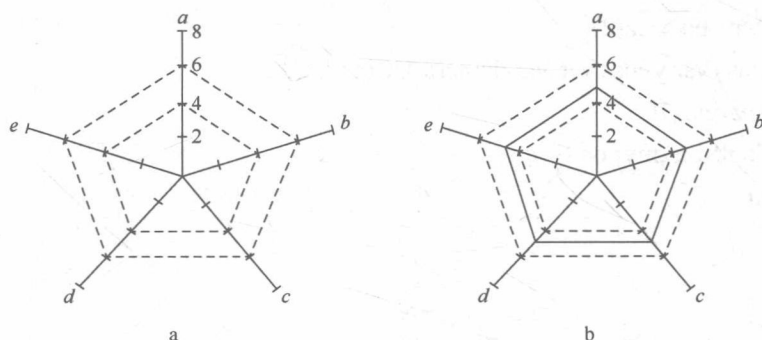


图 1-2 雷达图的使用

a. 标准雷达图; b. 表 1-4 中例 1 的  $(5 \pm 2)s$  雷达图

为了方便绘制雷达图, 笔者利用 MATLAB 编写了通用的绘图程序, 读者准备好必要的数后, 直接调用该绘图程序, 即可得到结果, 简洁方便, 下附 MATLAB 通用代码。

```
function [IsOk]=DrawRadar(RadarData,FactorName,SampleName,FileName)
```

```
% 用途: 用来绘制多元统计数据的雷达图。
```

```
% 数据: 包含 4 个输入参数, 其具体含义及使用如下:
```

```
% 1, RadarData 数据
```

```
% 按照如下格式准备 RadarData 数据, 比如调查学生的考试成绩, 分别是英语,
```

```
% 数学, 物理, 化学, 语文等, 则每列作为一个因子, 每行看作一个同学,
```

```
% 格式如下:
```

```
% % English, maths, physice, chemistry, chinese;
```

```
% 71.5 80.8 64.0 63.9 61.9 % Zhangsan %(张三)
```

```
% 69.2 60.8 69.0 64.0 63.8 % Lisi %(李四)
```

```
% 77.3 63.1 84.3 81.4 85.8 % Wangwu %(王五)
```

```
% 由于 matlab 兼容汉字的问题, 文字尽量使用英文标注。
```

```
% 2, 因子名称 FactorName 数据
```

```
% 为了输出图像, 还需要将各因素名称输入。本项数据属于可选项, 当没有本项数据时,
```



```

% 默认为 F1, F2, ..., 当有自己的准备数据时, 按照如下格式准备:
%   FactorName = {'Food', 'House', 'Doctor', 'Phone', 'Play'};
% 3, 样本点名称 SampleName 数据
% 在标注每个样本点数据时, 可采用文字名称或数字形式, 没有输入该项数据时,
% 自动按照 S1, S2, ..., 当采用准备好的文字名称标注时, 可按照如下准备 SampleName
数据
%   SampleName = {'Zhangsan', 'Lisi', 'Wangwu', 'Zhaoliu'};
% 4, 输出图片文件名称由 FileName 给定, 默认 RadarMap.jpg
% 内部参数:
% 1, 参数 OnOff 用来控制输出样本点的标注形式。当其值为 ON 时, 使用文字标注
% 当其值为 OFF 时, 使用数字标注。
% 2, 为了输出方便, 请给出输出图片的名称。
% 作者: mazhaipu, wdwsjlxx@163.com
%

% 1. 参数检测
if nargin<1||nargin>4
    error('输入参数数目错误! ');
end
if nargout>1
    error('输出参数太多! ');
end

% 2. 检查数据及格式
[nRows, nCols]=size(RadarData);
if nargin==1 %只输入 1 个参数时, 自动形成因子名称和样本点名称
    for jc=1:nCols
        FactorName{jc}=sprintf('F%d', jc);
    end
    for ir=1:nRows
        SampleName{ir}=sprintf('S%d', ir);
    end
    OnOff = 'OFF';
    FileName='RadarMap';
elseif nargin==2 %只输入 2 个参数时, 自动形成样本点名称
    if length(FactorName)~=nCols
        error('因子数目与数据不匹配! ');
    else
        for ir=1:nRows

```