多元分析方法及其应用

丁士晟 编著

吉林人民出版社

多元分析方法及其应用

丁士晟　编著

吉 林 人 民 出 版 社

多元分析方法及其应用

丁士晟　编著

*

吉林人民出版社出版　吉林省新华书店发行

长春新华印刷厂印刷

*

850×1168毫米32开本　24印张　530,000字

1981年12月第1版　1981年12月第1次印刷

印数：1—4,290册

统一书号：13091·92　定价：2.50元

内 容 简 介

本书介绍了多元分析方法及其应用，包括作者二十年来在这方面的工作成果，如用方差分析检验周期、用权重法建立 0.1 回归方程、点聚图的分析和检验。

全书共分十章，收集了一些较复杂和较新的多元分析方法，如聚类分析、模糊聚类、信息分析、第二类错误的检验、多因变量逐步回归、逐步多级判别。同时又从实用角度出发，选择了一些例子，由浅入深具体说明多元分析的方法，并详细列出了计算步骤，使读者通过范例所示，逐步掌握各种方法解决具体问题。为了便于读者应用，本书附录中有28个统计用表。

本书可供气象、农学、水文、地质、地震、医学等部门有关科技人员、研究人员以及大专院校师生参考。

前　　言

近几年来，多元分析在气象、农学、水文、地质、地震、医学等方面得到广泛应用，有必要写一本多元分析方法及其应用的书。本书试图尽量收集多元分析方法及其应用的最新成就，同时又从实用角度出发，选择了一些例子，由浅入深具体地说明多元分析方法，并详细地列出其计算步骤，使读者通过范例所示，能逐步掌握各种方法，解决具体问题。本书一共有 171 个例子，其中大部分是通例，虽然有一部分是气象方面的例子，但从事其它方面工作的同志也很容易理解。为了便于读者应用，本书附录中有28个统计用表，其中一部分是作者自己编算的，一部分录自 *Beyer, W. H.* 和山内二郎编著的统计用表，也有一部分录自常用数理统计用表等有关统计书。

全书共分十章，各章重点内容是：第一章叙述概率统计基本概念及数据处理；第二章介绍了各种统计检验，特别是第二类错误检验的方法和步骤；第三章是方差分析，从方差分析原理讲到多因素方差分析，同时为了科学安排试验，介绍了正交设计；第四章叙述各种回归分析方法，有复杂的多因变量逐步回归，也有简单有效的权重回归；第五章介绍了各种判别分析，有逐步多级判别分析和简便实用的点聚图的检验和分析；第六章介绍了信息分析及其应用；第七章叙述聚类分析，包括模糊聚类；第八章是滑动和插值；第九章讲的是拟合和配置实验曲线；第十章系统地介绍了过程分析，包括用方差分析方法检验周期。

此为试读，需要完整PDF请访问：www.ertongbook.com

李麦村同志审阅了本书，并提出了宝贵意见，吉林省气象科学研究所的领导和同志曾给作者鼓励和支持，在此一并致谢。由于作者水平有限，希望读者对本书的缺点和错误提出批评和指正。

丁 士 晟

1979年6月

目　　录

附　　录

第一章　概率统计基础

第一节　频率与概率

一　几个基本概念

自然界的现象可以分为两大类。一类是必然现象，这些现象符合因果规律，具有必然性。如在一个大气压情况下，温度低于0℃，水就要结冰，温度高于100℃，水就要沸腾。另一类是随机现象，在同样条件下重复试验或观测中，有时出现有时不出现具有偶然性的现象，又称为随机事件，简称为事件。如下雨同时打雷，就是随机事件，因为下雨时有时打雷有时不打雷。

在试验和观测中有若干随机事件 $A_1, A_2, \cdots A_i \cdots A_n$ 中至少发生某一事件所组成的事件 C，称为事件 C 是 $A_1, A_2, A_3, \cdots A_i \cdots A_n$ 诸事件之和。并记为

$$C = A_1 + A_2 + \cdots + A_i + \cdots + A_n = \sum_{i=1}^{n} A_i \qquad (1\cdot1)$$

这就是事件的和。

例1—1：如进行两次预报大风试验，A_1 表示只报对一次，A_2 表示两次均报对，A_3 表示两次均没有报对。

最多报对一次的事件 $C_1 = A_1 + A_3$

至少报对一次的事件 $C_2 = A_1 + A_2$

此为试读，需要完整PDF请访问：www.ertongbook.com

如果有一组事件 $A_1, A_2, A_3, \cdots A_i, \cdots A_n$ 同时发生的事件记为C，则事件C为$A_1, A_2, A_3, \cdots A_i \cdots A_n$诸事件之积，并记为：

$$C = A_1 \cdot A_2 \cdot A_3 \cdots A_i \cdots A_n = \prod_{i=1}^{n} A_i \qquad (1\cdot 2)$$

例 1 — 2：如果 A_1 表示下暴雨事件，A_2表示有洪峰事件，A_3 表示水库放流事件，C 事件表示下暴雨、有洪峰同时水库放流事件，则有

$$C = A_1 \cdot A_2 \cdot A_3$$

称C事件是事件 A_1、A_2、A_3 三个事件之积。

如果有 A、B 两个事件，并把不可能事件记为$\vee$，把必然事件记为$\sqcup$。

若$A \cdot B = \vee$，则事件A和事件B为互不相容事件。

若$A \cdot B = \vee$，同时有$A + B = \sqcup$则事件A和事件B称为互逆事件，事件A称为事件B的逆事件，事件B也称为事件A的逆事件。

并记为$\overline{A} = B \qquad \overline{B} = A$

下面介绍几个概率论、数理统计常用名词。

总体：由同性质成员或个体所构成的集体称为总体，它是我们研究对象的全体。总体所包含的成员往往是无穷的，是设想的或抽象的全体。总体又称为母体、母群、全群。

样本：由总体中抽出部分同性质的成员或个体所组成的集体称为样本。在实际工作中不可能得到总体的全部信息，只能用样本来推测总体的性质，处理实际问题往往和样本打交道。样本又称为子群、子样。

个体：总体或样本中某一个成员称为个体。每次观测或测量得到的数据就是一个个体的数据。

例 1 — 3：某河流解放后才有水文观测资料，由 1949 年到

1978年有30年流量资料，这30年水文资料就是样本。而某河流流量的总体是指有地球以来该河流的流量情况，既包括有资料的30年流量情况，也包括没有水文观测的历史时代的流量情况。而每次水文观测就是一个个体。

样本容量：由观测或试验得到的样本所包含个体的数量称为样本容量。一般统计上把样本容量大于30称为大样本，样本容量小于30称为小样本。

频数：某一样本在某一条件下出现的次数称为频数，记为M_i。

例1—4：如某森林测量一地块的树木直径，其直径≤10.0厘米的有60棵树，直径在10.1到20.0厘米的有80棵树，直径20.1到30.0厘米的有100棵树，直径≥30.1厘米的有50棵树。可以记为$d \leq 10.0$厘米时频数$M_1 = 60$，$10.1 \leq d \leq 20.0$时频数$M_2 = 80$，$20.1 \leq d \leq 30.0$时频数$M_3 = 100$，$d \geq 30.1$时频数$M_4 = 50$

频率：某一样本在某一条件下出现的次数与样本容量的比值称为频率，并记为f_i

$$f_i = \frac{M_i}{N} \qquad (1\cdot3)$$

其中N为样本容量，M_i为频数。频率是描述某一样本在某一条件下出现的可能性。

例1—5：据例1—4计算频率可以有

样本容量：$N = 60 + 80 + 100 + 50 = 290$

$d \leq 10.0$厘米的频率　$f_1 = \frac{M_1}{N} = \frac{60}{290} = 0.207$

$10.1 \leq d < 20.0$厘米的频率　$f_2 = \frac{M_2}{N} = \frac{80}{290} = 0.276$

$20.1 \leq d < 30.0$厘米的频率　$f_3 = \frac{M_3}{N} = 0.345$

30.0厘米$\leq d$的频率　$f_4 = \frac{M_4}{N} = 0.172$

频率是小于或等于1的正数。即

$$0 \leqslant f \leqslant 1 \tag{1·4}$$

频数和频率在气象中应用很多，在气表中就有风向频数和频率，不同雨量出现的频数和频率，不同天气现象出现的频数等等。

大量的试验证明，事件的频率具有一定稳定性。当样本容量很大时，频率在某一数值周围摆动，事件发生的可能性可以用一个数来表示。这个描述事件A在试验中发生可能性程度小于一的正数叫随机事件的概率，记作$P(A)$。

可以认为当样本容量$N\to\infty$时，事件的频率就是概率。

$$f(A)_{N\to\infty} = P(A) \tag{1·5}$$

当N很大时，频率$f(A)$是概率$P(A)$很好的近似值。

上面介绍了由频率出发来定义概率，下面再介绍一下概率的古典定义。假设试验的一切可能结果中，可以由N个互不相容事件，并且是等可能事件所组成的完备事件组，而其中有M个事件是有利随机事件A的，则随机事件A的概率等于有利于随机事件A出现的次数M与事件的总数N的比值。

$$P(A) = \frac{M}{N} \tag{1·6}$$

例1—6：如某地16个风的方位是等可能性的，求偏南风的概率。偏南风包括ESE，SE，SSE，S，SSW，SW，WSW，七个方位，偏南风的概率　$P = \frac{7}{16} = 0.4375$

如果我们在事件B已经发生条件下计算事件A的概率，则这种概率叫做事件A在事件B已发生的条件下的条件概率，记作$P(A|B)$。

例1—7：如果打雷的事件记为B，在打雷时下雨事件A的

概率为0.32，则这种概率为打雷情况下下雨的条件概率，记作 $P(A|B)=0.32$。

二　概率加法定理

概率加法定理：两互不相容事件的和的概率，等于这二事件的概率的和。事件A、事件B，若$A\cdot B=V$　则$P(A+B)=P(A)+P(B)$　　(1·7)

证明如下：

如有N个事件组成完备事件组，其中有M_1个有利事件A，M_2个有利事件B，并且事件A和事件B为互不相容事件，所以对A和B事件的有利事件共M_1+M_2个，于是有：

$$P(A+B)=\frac{M_1+M_2}{N}=\frac{M_1}{N}+\frac{M_2}{N}$$

$$=P(A)+P(B)$$

例1—8：如A、B是互不相容事件，事件A的概率$P(A)=0.21$，事件B的概率 $P(B)=0.12$，则这二事件的概率的和$P(A+B)=P(A)+P(B)=0.21+0.12=0.33$

同样对于多个互不相容事件的概率加法定理为：有限个互不相容事件之和的概率等于这些事件的概率的和。

有事件 A_1、A_2、A_3、……A_i……A_n为互不相容事件，$A_1\cdot A_2\cdot A_3\cdots\cdots A_i\cdots\cdots A_n=V$　　则

$$P(A_1+A_2+A_3+\cdots\cdots+A_i+\cdots\cdots+A_n)$$
$$=P(A_1)+P(A_2)+P(A_3)+\cdots\cdots+P(A_i)$$
$$+\cdots\cdots+P(A_n) \qquad (1\cdot 8)$$

对于事件 A_1、A_2、A_3、……A_i……A_n组成互不相容的完备事件组，即

$$P(A_1)+P(A_2)+P(A_3)+\cdots\cdots+P(A_i)$$

$+\cdots\cdots+P(A_n)=1$

$A_1\cdot A_2\cdot A_3\cdots\cdots A_i\cdots\cdots A_n=V$　　则

$$P(A_1+A_2+A_3+\cdots+A_i+\cdots+A_n)$$
$$=P(A_1)+P(A_2)+P(A_3)+\cdots+P(A_i)$$
$$+\cdots+P(A_n)=1 \qquad (1\cdot9)$$

对于互逆事件A和$\overline{A}$，因为有

$P(A)+P(\overline{A})=1$

$A\cdot\overline{A}=V$

则　$P(A+\overline{A})=P(A)+P(\overline{A})=1$

也即互逆事件的概率和等于1。

例1—9：如 A、B、C 是互不相容事件，事件A的概率$P(A)=0.21$，事件B的概率$P(B)=0.12$，事件C的概率$P(C)=0.23$。A、B、C这三事件的概率的和

$$P(A+B+C)=P(A)+P(B)+P(C)$$
$$=0.21+0.12+0.23=0.56$$

例1—10：如某人做水文预报20次报对15次，报错5次，那末预报3次有报错的概率是多少。

报错一次的概率　$P(A_1)=\dfrac{C_5^1C_{15}^2}{\sum\limits_{i=0}^{3}C_5^iC_{15}^{3-i}}=\dfrac{525}{1140}=0.46$

报错二次的概率　$P(A_2)=\dfrac{C_5^2C_{15}^1}{\sum\limits_{i=0}^{3}C_5^iC_{15}^{3-i}}=\dfrac{150}{1140}=0.13$

报错三次的概率　$P(A_3)=\dfrac{C_5^3C_{15}^0}{\sum\limits_{i=0}^{3}C_5^iC_{15}^{3-i}}=\dfrac{10}{1140}=0.01$

报错的概率：

$$P(A_1+A_2+A_3)=P(A_1)+P(A_2)+P(A_3)=0.60$$

三　概率乘法定理

概率乘法定理：二事件的积的概率等于其中一事件的概率与另一事件在前一事件已发生条件下的条件概率的乘积。

$$P(A\cdot B)=P(A)\cdot P(B|A)$$
$$=P(B)\cdot P(A|B) \tag{1·11}$$

证明如下：

如有N个事件组成完备事件组，其中有利于事件A的有M_1个，有利于事件B的有M_2个，而有利于事件$A\cdot B$的有M个，显然$M\leqslant M_1$，$M\leqslant M_2$。如果已知事件A发生，则表明有利于事件A的M_1个基本事件中必有一件发生。所以这时有利于事件B的基本事件仅有M个。

$$P(B|A)=\frac{M}{M_1}=\frac{\frac{M}{N}}{\frac{M_1}{N}}=\frac{P(AB)}{P(A)}$$

可得 $P(AB)=P(A)P(B|A)$

同理有 $P(AB)=P(B)P(A|B)$

同样可以推广到多个互不相容事件的概率乘法定理。

有限个事件的积的概率等于这些事件概率的乘积，其中每一事件的概率是在它前面一切事件都已发生条件下的条件概率。

$$P(A_1A_2A_3\cdots A_i\cdots A_n)$$
$$=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots\cdots$$
$$P(A_n|A_1A_2\cdots A_{n-1}) \tag{1·12}$$

例　1—11：某地7、8月打雷的概率为0.2，下暴雨的概率为0.05，下暴雨打雷的概率为0.02，求打雷条件时下暴雨的概率和

此为试读，需要完整PDF请访问：www.ertongbook.com