

57.15
420

地貌统计

黄镇国
方国祥
宗永强
林晓东
黄光庆
编著



科学普及出版社广州分社

地貌统计

宗永强 林晓东 黄光庆

科学普及出版社广州分社

内 容 简 介

本书以较多的实例介绍常用统计方法在地貌学研究中的应用，内容包括定量描述或揭示地貌现象的量的分布特征、异同程度、依存关系、空间分布特征，预测与控制、分类与分区等、实用性强，深入浅出，可供从事地貌学、地质学、地理学及相邻学科研究的读者参考应用。

地 貌 统 计

黄镇国等

*

科学普及出版社广州分社出版发行

广东省新华书店经销

广地印刷厂印刷

*

开本：787×1092毫米 1/32 印张：9.4 字数：180千字

1989年10月第1版 1989年10月第1次印刷

印数 1—1000册

ISBN7-110—1327-8/K.26

定价：4.50元

前 言

长期以来，地貌学研究大多是进行定性的调查分析，然后经过归纳综合，得出某种地貌现象的规律性。但是，这种研究方法对地貌内在规律的揭示是不够充分的，因而在一定程度上限制了地貌学的发展。

自本世纪 50 年代末至 60 年代中以来，地貌的量化技术或计量地貌学有了长足的进展。地貌的制约因素很多，而且彼此之间错综复杂，所以现阶段计量地貌学逐渐趋向用多变量（多元）统计分析的方法。变量多，计算工作量就很大，利用常规的计算手段，势难胜任。但是，自本世纪 40 年代第一台电子计算机问世以来，电子计算机就成了计量地貌学最有效的运算工具，从而促进了地貌学研究进一步向量化发展。

计量地貌学的内容很多，主要有以下几个方面。

定量描述地貌要素的量的分布特征：例如相对数、平均数、标准差、变异系数、分异度等。定量分析地貌现象的异同程度：例如显著性检验、方差分析等。定量揭示地貌要素之间的依存关系：例如相关分析等。定量表述地貌要素的空间分布特征：例如趋势面分析。对地貌现象进行预测：例如置信限、回归分析等。对地貌现象进行定量的分类或分区：例如因子分析、对应分析、判别分析、聚类分析等。

地貌学研究的另一个发展趋向是越来越重视积累野外的或室内的观测实验数据，用以阐明地貌现象的本质特征，而

如何有效而充分地处理大量的数据，导出规律性的结论，这比积累数据更加重要。因此，常用的数理统计方法是地貌学研究中不可缺少的手段。

鉴于上述理由，我们根据多年来在科研实践中较广泛地应用地貌统计方法的经验，选择较多的实例，并参考有关数理统计的书籍，编写了这本介绍地貌学常用统计方法的书。

本书的特点除举出较多的实例（约70个）之外，还重视通俗性，力求深入浅出地介绍各种统计方法的演算过程，而较少介绍数学原理和公式的推导，俾使读者易于掌握这些方法，于实践中运用。

必须指出，本书并非一本理论地貌学的著作，所以其内容不可能涉及地貌学研究中的一系列物理的或数学的方程。关于地貌统计方法，也是择其要者，加以介绍。

本书的编写分工如下：第一章至第三章由黄镇国执笔，第四章和第五章由黄镇国、林晓东执笔，第六章由黄镇国、宗永强执笔，第七章至第九章由方国祥执笔，第十章由宗永强执笔，第十一章由宗永强、黄光庆执笔。全书由黄镇国修改定稿。

由于编写者的实践经验不多，地貌学和数学的知识也很有限，书中一定存在不妥的甚至错误的地方，敬希读者批评指正。

目 录

第一章 相对数	(1)
§ 1-1 相对数与绝对数	(1)
§ 1-2 比	(1)
§ 1-3 率	(3)
§ 1-4 率的标准化	(8)
1. 另选一个共同标准	(9)
2. 以两组资料之和作为共同标准	(9)
3. 任选一组资料作为共同标准	(10)
§ 1-5 相对比	(12)
1. 对比指标	(13)
2. 相关指标	(13)
3. 动态数列	(13)
第二章 平均数	(15)
§ 2-1 算术平均数	(15)
1. 小样本资料的平均数	(15)
2. 大样本资料的平均数	(16)
§ 2-2 中位数	(22)
1. 小样本资料中位数的计算法	(22)
2. 大样本资料中位数的计算法	(22)
§ 2-3 众数	(24)
§ 2-4 百分位数	(25)
§ 2-5 几何平均数	(27)
第三章 标准差	(30)
§ 3-1 全距	(30)
§ 3-2 离差	(31)
§ 3-3 方差	(32)

§ 3-4	标准差	(34)
1.	小样本资料的标准差算法	(35)
2.	大样本资料的标准差算法	(38)
§ 3-5	就异系数	(40)
第四章	分异度	(42)
§ 4-1	简单分异度	(43)
§ 4-2	复合分异度	(43)
1.	信息函数	(43)
2.	均衡度	(46)
§ 4-3	分异度与沉积环境	(49)
第五章	显著性检验	(55)
§ 5-1	总体与样本	(55)
§ 5-2	抽样误差的表示方法 (标准误)	(56)
1.	平均数标准误 (SE) 的计算	(57)
2.	相对数 (率或比) 标准误 (Sp) 的计算	(58)
§ 5-3	总体平均数或相对数的置信限	(59)
1.	总体平均数的置信限	(60)
2.	总体相对数的置信限	(60)
3.	用 t 值估测总体的置信限	(61)
§ 5-4	平均数的显著性检验	(62)
1.	小样本平均数的显著性检验 (t 检验)	(63)
2.	大样本平均数的显著性检验 (u 检验)	(68)
3.	非正态分布平均数的显著性检验	(69)
§ 5-5	相对数的显著性检验	(75)
1.	样本率 (比) 与总体率 (比) 的差异显著性检验	(75)
2.	两个样本率 (比) 的差异显著性检验	(76)
3.	卡方 (χ^2) 检验	(77)
4.	行列表的 χ^2 检验	(79)
5.	四格表法	(80)
第六章	方差分析	(82)

§ 6-1	单因子方差分析	(83)
§ 6-2	双因子方差分析	(88)
§ 6-3	交互作用的方差分析	(91)
§ 6-4	多因子方差分析	(95)
第七章	因子分析	(100)
§ 7-1	单因子分析	(100)
§ 7-2	双因子分析	(100)
§ 7-3	多因子分析	(104)
1.	多因子分析的步骤	(104)
2.	多因子分析的计算机程序	(110)
§ 7-4	对应分析	(119)
1.	对应分析的步骤	(119)
2.	对应分析的计算机程序	(122)
第八章	相关分析	(128)
§ 8-1	相关的种类	(128)
1.	单相关	(128)
2.	复相关	(129)
3.	偏相关	(129)
§ 8-2	相关系数的计算	(129)
1.	小样本相关系数的计算	(130)
2.	大样本相关系数的计算	(132)
第九章	回归分析	(134)
§ 9-1	一元线性回归	(134)
1.	一元线性回归方程的建立	(134)
2.	一元线性回归方程的显著性检验	(136)
3.	一元线性回归方程的估测误差	(136)
4.	一元线性回归方程的预测	(138)
§ 9-2	多元线性回归	(138)
1.	多元线性回归方程的建立	(138)
2.	多元线性回归方程的方差检验	(144)

3. 各个自变量回归效果的检验	(145)
4. 多元线性回归的计算机程序	(149)
§ 9-3 可以化为一元线性回归问题的非线性关系	(156)
§ 9-4 多元非线性回归	(159)
§ 9-5 逐步回归分析	(163)
1. 逐步回归的准备	(163)
2. 逐步回归的步骤	(164)
3. 逐步回归的计算机程序	(178)
§ 9-6 趋势面分析	(187)
1. 多项式趋势面的计算	(187)
2. 趋势面拟合效果的检验	(191)
3. 趋势面分析结果的解释	(192)
4. 趋势面的计算机分析程序	(195)
第十章 判别分析	(202)
§ 10-1 二级判别	(202)
§ 10-2 逐步判别	(206)
第十一章 聚类分析	(216)
§ 11-1 数据的平均化处理	(218)
§ 11-2 数据的标准化处理	(221)
§ 11-3 求距离	(223)
§ 11-4 样本聚类	(226)
1. 最短距离法	(226)
2. 最长距离法	(229)
§ 11-5 相似系数聚类	(233)
§ 11-6 有序样品聚类	(239)
§ 11-7 模糊聚类	(240)
1. 模糊聚类的步骤	(240)
2. 模糊聚类的计算机程序	(245)
附表 1-11	(248)

第一章 相对数

§ 1-1 相对数与绝对数

说明研究对象的原始实际数据，称为绝对数。例如在地貌学研究中搜集到的原始数字资料：A区有400条岩石节理，其中北东向的有13条 B区有180条岩石节理，其中北东向的有6条；对10种相同的地貌类型，测量其坡长，得出10个测量数值为93, 102, 110, 102, 98, 109, 92, 97, 100, 103m，这些数值都是绝对数。从这些绝对数中，往往不容易看出它们的意义和规律性。通过一定的统计处理，可以把这些原始数据加以概括提炼，使之成为简明的可以用来对地貌现象的规律性进行适当估计的数值，即把有关系的绝对数建立一定的比例关系，这个比例关系称为相对数或统计指标。

绝对数和相对数各自说明一个问题，不能混淆。但两者能相互补充综合地反映问题的性质。例如某地的海滩面积，1984年为6.35万亩，比1895年增长2.8%。前者(6.35万亩)为绝对数，反映海滩的面积；后者(2.8%)为相对数，反映1984年比1895年海滩增加了多少，说明一种关系。

常用的相对数有比、率、相对比。

§ 1-2 比

“比”是一种用来表示事物或现象内部构成情况的统计指

标，又称构成比、构成指标、结构指标，说明事物内部某一组成部分的比重，常用百分比来表示。计算公式为：

$$\text{比} = \frac{\text{某一构成部分的个体数}}{\text{各构成部分个体数的总和}} \times 100\% \quad (1-1)$$

可以简单地写成：

$$\text{比} = \frac{\text{部 分}}{\text{全 体}} \times 100\%$$

“比”的特点是：“分子”是部分，“分母”是全体。部分小于全体，分子必小于分母。

【例 1-1】 某区域的土地总面积为725km²，其中台地为107km²，

$$\text{比} = \frac{107}{725} \times 100\% = 14.76\%$$

即该区域的各种地貌类型之中，台地占14.76%。

【例1-2】 如表1-1，在某花岗岩台地上，记录了不同类型侵蚀沟的条数，从百分比可以了解到各类侵蚀沟在全部侵蚀沟中所占的比重。例如细沟只占0.5%，而崩岗却占62.0%。由此可见此台地的侵蚀程度已十分深刻。

表1-1 某花岗岩台地上侵蚀沟条数统计

侵蚀沟类型	条 数	百分比(%)
细 沟	4	0.5
浅 沟	26	3.5
切 沟	38	5.0
冲 沟	217	29.0
崩 岗	465	62.0
合 计	750	100.0

由表1-1可见，比的总和必等于100。若某一部分的比减少，其他部分的比必然相应增加。各分组的比是可以相加的。

§1-3 率

绝对数和相对数在统计学上还表现为频数和频率。频数是指在做观测或实验时，每一种情况出现的次数，属于绝对数。频数占总次数的比率称为频率，属于相对数。频率又称率或频率指标。因此，率是表示在一定条件下某现象在全体中所出现的频繁程度。常以百分率(%)、千分率(‰)、万分率(1/万)、十万分率(1/10万)等表示。计算公式为：

$$\text{率} = \frac{\text{某种现象发生的例数}}{\text{可能发生该现象的主体个数}} \times 100\% (\text{或} 1000\% \dots) \quad (1-2)$$

可以简单地写成：

$$\text{率} = \frac{\text{发生例数}}{\text{主体个数}} \times 100\% (\text{或} 1000\% \dots)$$

【例1-3】 如§1-1所举出的绝对数：A区有400条岩石节理，其中北东向的有13条。13为发生例数；400为主体个数，即这400条节理都存在表现为北东向的可能性。A区和B区北东向节理的发生率可计算如下：

$$\text{A区北东向节理率} = \frac{13}{400} \times 100\% = 3.3\%$$

$$\text{B区北东向节理率} = \frac{6}{180} \times 100\% = 3.3\%$$

根据这两个频率指标可以看出，A区和B区的北东向节理的频率是一样的，可能反映两区的岩性、构造方向是很相似的。

【例1-4】为了分析某区域断裂构造方向与河流走向的关系，将主要河流的流路由曲线简化为折线，绘成水系折线图，然后在地图上量测每条线段的方位角和长度(km)，得知共549条，总长806.8km，24个方位的河道长度如表1-2所示，从中可以求得24个方位的河流流向频率(%)。

表1-2 某区域河流流向频率(%)

方 位	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180°
河流长度(km)	72.6	67.7	65.3	66.9	48.4	32.2	66.1	50.0	80.6	83.1	66.1	65.3
频 率	9.0	8.4	8.1	8.3	6.0	4.0	8.2	6.2	10.0	10.3	8.2	8.1
方 位	195°	210°	225°	240°	255°	270°	285°	300°	315°	330°	345°	360°
河流长度(km)	75.0	64.5	66.1	64.5	50.8	33.8	64.5	40.3	82.2	96.8	64.5	66.9
频 率	9.3	8.0	8.2	8.0	6.3	4.2	8.0	5.0	10.2	12.0	8.0	8.3

由24个方位的河流流向频率，可绘成河流走向频率玫瑰图，由这种图可以看出，本区河流的优势方向为北北西—南南东，次优势方向为北北东—南南西，说明北东向和北西向两组断裂构造对本区河流的发育都有影响。由优势的构造方向，可进一步分析构造应力场。

上述例子中的“比”和“率”，其计算方法很简单，而且都用百分(%)表示，但是，“比”和“率”所表达的意义并不相同，使用时必须注意，否则将导致错误的结论。

“比”与“率”的区别在于：(1)“比”是表达部分与整体的

关系；“率”是表达发生例数与主体个数的关系，分子是分母发生的现象。(2)“比”是可以相加的，其总和必等于100；“率”是不能相加的，其总和不等于100。(3)“比”的分子必小于分母；“率”的分子可以小于、等于或大于分母。为了说明“比”和“率”的区别，下面举一个例子。

【例1-5】 如表1-3，某台地的总面积为 167.5m^2 ，按坡度等级分别量算了面积及侵蚀沟条数。现在要了解哪一个坡度等级的侵蚀程度最严重。从第(4)栏的侵蚀比来看， $6-12^\circ$ 的坡面为46.66%， $>20^\circ$ 的坡面为35.56%，前者大于后者，但不能由此得出结论，在岩性、外力条件等都很相似的小区

表1-3 各坡度组的坡面侵蚀程度

(1) 坡度等级	(2) 面积(m^2)	(3) 侵蚀沟条数	(4) 各坡度组的 侵蚀比(%)	(5) 各坡度组的 侵蚀率(%)
3— 6°	82.9	4	4.44	4.8
6— 12°	28.7	42	46.67	146.3
12— 20°	46.6	12	13.33	25.7
$>20^\circ$	9.3	32	35.56	344.0
合 计	167.5	90	100.00	53.7

域内， $6-12^\circ$ 的坡面反而比 $>20^\circ$ 的坡面侵蚀严重。原因在于第(4)栏的比值是表达各坡度等级的侵蚀沟条数占该台地侵蚀沟总条数的百分比，例如 $6-12^\circ$ 坡面的侵蚀沟为42条，占台地侵蚀沟总条数(90)的46.66%。可是各坡度等级的面积很不相同，要了解侵蚀程度，必须用“率”来表达，而不能

用“比”来表达。以第(3)栏的侵蚀沟条数为分子，以第(2)栏的面积为分母，求出第(5)栏的侵蚀率，表示每 m^2 面积内有多少条侵蚀沟，才能说明侵蚀程度。显然， $>20^\circ$ 坡面的侵蚀率为344%，是最严重的。从表1-3还可看出，“比”是可以相加的，其总和必等于100。“率”是不能相加的，不能说该台地的侵蚀率是第(5)栏4个值的总和(520.8%)，而应以第(3)栏的合计值(90)为分子，以第(2)栏的合计值(167.5)为分母，求得该台地的侵蚀率为53.7%。从表1-3还可看出，“比”的计算，分子必小于分母；“率”的计算，分子可以小于分母，也可以大于分母。

为了进一步说明“比”的应用，根据“比”的总和必等于100的特点，下面再举两个例子，看一看怎样才是正确的结论。

【例1-6】如表1-4，在一个柱状剖面中，采取5层不同岩性的沉积物样品，分别检出化石硅藻的个体数，求出各自的百分比，其中粘土的硅藻个数179个，占42.5%。但不能由此得出结论，粘土的硅藻含量最丰富，因为没有给出提供对照的条件例如沉积物样品的体积、各个样品的沉积环境是否一致。

表1-4 沉积物岩性与化石硅藻含量的关系

岩 性	化石硅藻个数	百 分 比
粘 土	179	42.5
粉 砂	92	21.8
细 砂	75	17.9
中 砂	32	7.6
砂 砾	43	10.2
合 计	421	100.0

【例1-7】如表1-5，根据同一地区的1962年和1983年的1/5万地形图查得三类侵蚀沟的条数，并求出百分比。其中崩岗由于进行了治理，从1962年的74条减为30条，冲沟和切沟的条数也有减少。如果从“比”来进行比较，就会得出冲沟和切沟条数增加的错误结论。原因在于“比”的总和必等于100，某一部分的比减少时，其他部分的比就相应增多。崩岗的比由71.2%减为53.6%，而比的总和要等于100，就影响到冲沟和切沟的比重相应增加，并不是后两者的条数增加了。

表1-5 某地区两年侵蚀沟条数的百分比

侵蚀沟类型	1962年		1983年	
	条 数	%	条 数	%
崩 岗	74	71.2	30	53.6
冲 沟	13	12.5	12	21.4
切 沟	17	16.3	14	25.0
合 计	104	100.0	56	100.0

值得指出，虽然“率”一般可作比较，但决定“率”的因素往往是多方面的，有影响的因素都应相同或很接近，才能进行“率”的比较。还要注意样本例数是否足够。例数过少时，用绝对数表示较好。两个“率”相比较是否有显著差别的问题，将在第五章§5-5叙述，即“率”的比较，不能只凭表面数值的大小来作结论，而应进行显著性检验。

同一命题内部两种不同的“率”进行比较时，要加以标准化。

§1-4 率的标准化

【例1-8】如表1-6，A、B两个地点，分别测量了100个砾石，砾石的岩性有脉石英、变质岩、火山岩三类。两个地点所观察的三类砾石的个数不相同，A处为60，30，10个；B处则相反，为10，30，60个。其中属于第5级磨圆度的砾石个数也不相同，A处为45，25，9个；B处为7，24，54个。分别求得第5级磨圆度砾石的出现频率。从三类砾石各自的频率来比较，A处都比B处高，表明前者砾石的磨圆程度比后者好。但是，从“总率”来看，却出现了矛盾现象，后者(85%)比前者(80%)高。究竟哪个地点的砾石磨圆程度较好呢？

表1-6 两个地点砾石磨圆状况的比较

砾石岩性	A 地 点			B 地 点		
	砾石个数	第5级磨圆度的砾石个数	频率 (%)	砾石个数	第5级磨圆度的砾石个数	频率 (%)
脉石英	60	45	75	10	7	70
变质岩	30	25	83	30	24	80
火山岩	10	9	90	60	54	90
合 计	100	79	79	100	85	85

在这种情况下，要用统计学的“标准化法”对两组资料进行校正后，再作比较。率的标准化有三种方法，现分述如下：