

# 关于岩土样本统计分析的几点讨论

李征翼

(中基发展建设工程有限责任公司, 河北三河 065201)

**【摘要】** 结合岩土单元体的天然成因及人为设定的基本特性, 阐述了岩土样本离散的必然性及其离散程度与岩土单元体分划范围的相关性, 通过对变异系数的适用性及岩土样本概况性的分析, 提出采用均同指数评价样本分布的离散程度和按相对容量标准控制岩土样本数量等观点, 建议切实针对岩土母体的分布形态推断标准值, 同时对小样本条件下标准值的确定方式进行了分析和建议。

**【关键词】** 岩土工程; 样本; 统计分析; 离散程度; 样本容量; 置信水平

**【中图分类号】** TU 192

## Viewpoints on Statistical Inference for Geotechnical Sampling

Li Zhengyi

(China Solibase Engineering CO., LTD, Sanhe Hebei 065201 China)

**【Abstract】** It is common that geotechnical samples are delimited according to the natural geological origin and period. The necessity of dispersion of geotechnical sampling probability distribution, and the degree of diversity is relative to partitioned range of the base layer are expounded. Based on the analysis result of the applicability of variation coefficient and the generality of sampling, the use of the resemblance index to appraise the degree of the dispersion is advanced, and the number of geotechnical samples must be strictly controlled according to the principle of relative capacity. The author suggests, according to the essence of the base layer to infer the risk limit data, should be by another method in a small amount samples condition.

**【Key Words】** geotechnical engineering; sampling; statistical inference; degree of dispersion; capacity of the sample; probability level

### 0 引言

可靠性分析理论应用于岩土工程勘察技术工作是科学发展的需要。尽管现行规范已规定了岩土样本统计分析的方式和内容, 但在实际工作中, 技术人员对分析项目及相关要求仍持有不同程度的认识。如何科学运用数理统计理论, 正确理解岩土样本统计分析工作的实质, 以及合理选用岩土工程特性参数, 依然是值得进一步探讨的问题。

#### 1 样本的离散特性

变异系数是目前通用的评价岩土样本离散性的统计分析指标, 至今一些标准中, 仍然保留着依据变异系数划分岩土样本离散程度等级的参照标准和不同岩土参数变异系数的经验值。

变异系数  $\delta$  由统计推断出的总体分布特征参数<sup>[1]</sup> 定义和计算:

$$\delta = \frac{s}{\mu} \quad (1)$$

式中:  $\mu$  为平均值;  $s$  为样本标准差。

相应计算公式<sup>[2]</sup> 如下:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2} \quad (3)$$

式中:  $n$  为子样的样本容量, 即通俗意义上的样本数量。

引设变异系数, 其中的主要原由是原本评价离散特性的指标  $s$  带有量纲, 不易于形象化理解和比较不同性质概率分布的离散形态<sup>[3]</sup>。

通过定义公式(1)可以看出, 变异系数的本质是所推断的总体散度特征参数  $s$  相对于自身分布位置

特征参数  $\mu$  的比值, 属于相对性指标, 因此将其作为绝对性参数检测不同母体的离散程度, 则会因为位置特征的差别, 出现不公正的判断和认识:

两组具有相同散度特征参数值的概率密度分布示例见图 1, 根据变异系数计算公式可得出如下的评价结论:

$$\begin{aligned} & \text{因为} && s_1 = s_2 \\ & \text{及} && \mu_2 > \mu_1 \\ & \text{所以} && \hat{\delta} = \frac{s_1}{\mu_1} > \hat{\delta} = \frac{s_2}{\mu_2} \end{aligned}$$

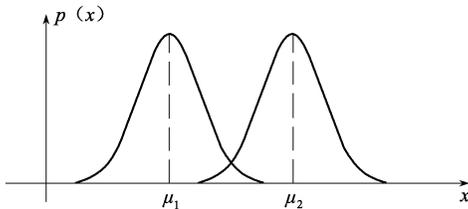


图 1 两组相同离散程度的正态分布概率密度曲线

示例分析证明变异系数指标确实未能确切地反映概率分布的离散程度, 而是随着位置特征的差异, “夸大”或“低估”了散度特征的真实性和在位置特征值为零的特殊条件下,  $\delta$  值凸现为无穷大, 变异系数失去了意义。

事实上, 概率分布散度特征表现的相关形态不同(见图 2), 离散程度大的密度曲线形态宽阔和矮, 反之则呈现窄小而高耸。可见显现曲线“峰顶”高低的众数概率密度值同时亦能反映分布函数的离散特征, 其数值与离散程度呈反比关系。姑且将离散程度的逆向特征表述为概率分布的均同性。

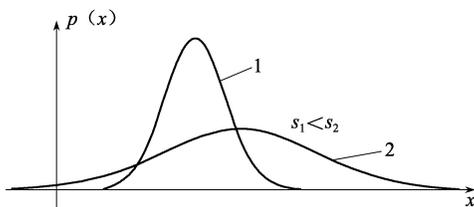


图 2 不同离散程度的正态分布概率密度曲线

根据正态分布模型概率密度函数公式:

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right]$$

当  $x = \mu$  时, 由统计推断的总体概率分布函数得知其众数值(平均值)所对应的概率密度值为

$$p(\mu) = \frac{1}{s \sqrt{2\pi}} = \frac{0.399}{s} \quad (4)$$

综上所述, 本文试定义均同指数  $a$  作为描述岩土样本概率分布离散特性的指标:

$$a = \frac{50.0}{s} \quad (5)$$

均同指数形式上只与散度特征参数相关, 实质是将经统计推断的正态分布特征概化为  $[\mu - s, \mu + s]$  区间范围内均匀分布模型条件下所对应的概率密度值的百分数。 $a$  值越高, 表明其离散程度越小, 即占有的样本值均等或相同。由于摒除了位置特征参数(平均值)的干扰, 因此这一指标适用于确切地评价不同批次子样间的离散特性。与传统的散度特征标准差相比, 均同指数将离散特性参数值同化为以概率密度为基础的指数形式, 以便于实际应用中的形象化理解或比较不同量纲条件下样本的离散条件。

众所周知, 岩土体是受内外地质作用综合影响形成的自然结果。必定与按人为设定标准控制加工的产品不同, 天然堆积土层无论水平方向, 还是垂直方向上的相变或不连续是必然的, 即不均匀性是岩土样本母体的固有形态, 并且不同成因的岩土体的离散程度亦不尽相同。另外, 由于系统误差性质的差异, 岩土样本值的散度特征同时与其试验类型和方法相关, 即同一母体不同岩土参数的均同性亦不可能按相同标准评鉴。以往建议的变异系数经验性参考值通常按地域分类限定, 即是岩土母体成因环境与样本数据类别两个离散特性影响因素的共同体。

同时必须注意的是, 基于岩土体离散的必然性, 样本母体的分划条件也将影响相关岩土参数的离散程度, 岩土单元体所涵盖的范围越大, 子样的离散程度就可能越高。因此, 实际技术工作中, 用于考评岩土样本离散程度时, 所参考经验值的范围性是难免的。

## 2 样本数量及其代表性

6 件样本的数量要求<sup>[4]</sup> 几乎是现行岩土技术规范的一致规定, 同时又强调当土质不均匀时, 应增加样本数量。显然 6 件样本属于适用于规范规定的统计分析方法低限容量要求的单一性原则, 但这一绝对数量的规定, 确实为相应规范和标准的执行带来很多困惑。

为满足统计分析可靠性目标要求, 勘察技术质量工作的指导原则应是尽可能取得足够数量的代表性岩土样本。样本容量的大小反映了岩土工程师对客观事物了解的详尽程度, 较好地统计估计需要充分大的样本容量满足对总体特征的概括和认识, 否则可能导致统计拟合失真, 甚至产生与实际不符的错误分析结论。

数理统计理论将 30 个以上的样本容量称为大

样本,否则属于小样本;并认为50件以上甚至达到100件时,方才符合充分大的标准,显然在实际的岩土工程勘察工作中,难以每每实现这样的目标,而探讨岩土样本容量满足概括性要求的最低限制标准更具有其实际意义。

按概率计算方法确定的最少岩土样本数量与指标的变异系数及容许相对误差和规定的置信水平有关<sup>[5-9]</sup>,均同性好的母体可以要求较低的样本容量,反之则需要加大样本容量以求深入了解。但如在实际技术工作中,不依据岩土母体的实际条件,仅以单一性的绝对数量要求作为遵从标准,必然会出现样本数量难以掌控的现象。

为采取特定空间范围内分布母体的代表性样本,并满足可靠性分析目标的概括要求,符合统计理论的随机性抽样方法是规定在岩土体分布范围内按一定的间隔采集样本,而相应的样本抽取率必定与取样间隔相关。可见,科学合理的岩土样本数量要求应是结合适宜的抽取率,由统计分区范围内岩土单元体(母体)的量值所决定。相应的抽取率标准须结合岩土单元体对建设工程的影响程度及其复杂程度和相应建设工程的重要性等级等条件综合确定。

为体现占有样本数量对所代表岩土母体的概括程度,笔者试借鉴矿产地质中的“储量”概念,用以表象岩土母体的量值,同时将采集的岩土样本数量与其所代表储量(体积)的比值称之为岩土样本的相对容量,以此等同于普遍意义的样本抽取率。

相对容量对应的是满足特定“储量”岩土母体统计估计精度要求的代表性样本数量,其本质除体现传统统计理论控制样本抽取率的概括性要求外,同时突出强调了岩土样本统计分析限于所代表岩土单元体的针对性原则,并且为勘察工作质量控制提供考核依据。

依据相对容量标准限定岩土样本数量,可避免在大范围、多勘探点的勘察工作中随意缩减样本容量和舍弃必要样本的采取,并且能够制约只顾响应单一的数量要求,不注重岩土样本的针对性,将特定母体范围以外的非代表性样本肆意划归为子样的迷失性统计分析。

事实上,现行相关技术标准在规定了勘探点间距和采样点竖向分布标准<sup>[7]</sup>的同时,就已初步计划出了勘察评价范围内各岩土单元所能够采集的样本数量,也即是相对容量原则的具体体现。

因此,相关技术标准在制定勘探点间距的基础上,应进一步限定岩土样本采取的水平间隔和沿深

度方向的采样间距,而对样本绝对数量不作不切实际的规定。实际工作中,可根据相应岩土层的勘探累计进尺与所包含的岩土样本数量,简化相对容量的检验和评判。

### 3 统计推断与置信标准

现行工程技术规范中岩土样本统计分析的置信标准一般统一规定为95%,并结合岩土工程参数的安全储备原则,采取单侧置信界限值推断相关标准值 $\phi_k$ :

$$\phi_k = \gamma_s \phi_m \quad (6)$$

$$\gamma_s = 1 \pm \frac{t_\alpha}{\sqrt{n}} \delta \quad (7)$$

式中: $\phi_m$ 为平均值; $t_\alpha$ 为学生氏分布模型函数界限值; $\gamma_s$ 为统计修正系数; $\delta$ 为变异系数。

文献[4]将公式(7)中的 $\frac{t_\alpha}{\sqrt{n}}$ 项通过拟合成为以

下的近似公式:

$$\frac{t_\alpha}{\sqrt{n}} = \left\{ \frac{1.704}{\sqrt{n}} + \frac{4.678}{n^2} \right\} \quad (8)$$

同时在相关条文说明中,解释到公式(7)系依据随机场理论折减,并采用近似计算而确定的形式。但无论如何,现行标准值计算公式与统计理论,对只与系统误差影响相关“真值”的估计形式相同,其计算结果是参数母体平均值的无偏估计量;同时尽管采用学生氏分布模型修正计算标准值能够进一步增强小样本容量条件下岩土样本统计计算结果的可靠程度,但毕竟造成分析和评价前后选用数学模型的不一致。

鉴于岩土样本的随机分布不仅是由于技术工作的系统误差影响所致,因此,一定置信水平的标准值还应切实结合岩土母体所固有的分布特征实施统计分析。若依据平均值偶然性的置信估计作为可靠性标准,事实上是忽略了天然成因的岩土体本身客观存在离散性的这一基本特性。

另外,鉴于正态分布尾部小概率估计失真的模型自身缺陷等原因,建议采用90%的置信水平进行标准值修正计算:

$$\phi_k = \phi_m \pm 1.282s \quad (9)$$

式中: $s$ 为依据正态分布模型统计推断出的样本标准差。

关于岩土参数标准值的分析建议,尚有如以下几个问题须引起我们的注意:

1) 统计分析母体应针对建设工程影响范围及岩土样本的离散等级等条件综合划分。目前,大范

围场地、多项建设工程的勘察项目很多,一次涉及建设工程数十、甚至逾百,勘探点位成百上千的同期勘察项目经常遇到。如此状况下如果仍是按通体大范围划分的稳定分布岩土层作为统一的岩土单元,并且笼统按95%的置信标准分析推断岩土设计参数,那么在理论上按最不利条件组合,就有5%建设工程的风险概率可能处于失效范围内,同时,还有相当数量的建设工程则可能得到了不必要高的安全储备。显然,这样的推断结果已失去了统计分析的积极意义。

同理,初步勘察阶段在无法完全依据具体建设工程划分岩土单元体范围的情况下,如未按限定的样本离散等级划分统计区域或地段进行针对性地分析评价,则会出现统计推断结论含混,甚至致使所建议标准值的置信水平迷失。

2) 标准值修正的侧限关系应结合推断分析的目的,按不利组合的原则设置。如:通常沉降变形分析估计的是最大的可能变形值,压缩变形指标标准值计算的统计修正系数通常采用负号。但对于高低层相连建筑的差异沉降分析,如期望了解毗连建筑之间可能的最大沉降差,供设计或施工预留变形调节量时参考,此时需估计低层裙楼(房)可能出现的最小沉降量,则应要求统计修正系数改用正号计算来推断相应压缩变形指标的标准值。

3) 小样本容量条件下标准值,应结合工程重要性等级及相应的岩土工程背景条件,确定修正或取值方法。分析表明(见表1),式(8)等号两边的拟合结果随样本数量的逐步增加,在数量为6件时,规范推荐公式的相对误差由负值转折为正,所计算的标准值由此以后不再有突破理论风险界限的忧虑,因此6件样本确实可以坦然地作为该公式的适用界限。

表1 标准值近似计算公式误差分析表

样本数 $n$ /件	不同方法计算的界值		拟合结果 相对误差 /%
	学生氏分布查表	规范建议公式	
2	4 464.7	2 374.4	-46.82
3	1 685.8	1 503.5	-10.81
4	1 176.5	1 144.4	-2.73
5	0 953.5	0 949.2	-0.45
6	0 822.6	0 825.6	0.365

该近似公式在样本数量小于4时的拟合误差较大,毫无疑问相应标准值的推断计算须另辟蹊径;但对于4~5件样本拟合相对误差仅为-2.73%~

-0.45%精度条件下,还需岩土工程师结合样本的均同性等条件决断该公式的适宜性。

部分现行标准建议在仅取得1件样本时的岩土参数标准值按固定系数折减计算。但如果某一岩土样本代表的是一个相当小储量的岩土体,则其可能已具有较高的概括性,所以小样本条件下岩土参数标准值的建议计算不可依据经验一概而论。

#### 4 结论及建议

变异系数反映的是概率分布散度特征与位置特征的相对关系,不能确切反映样本分布的离散程度,且在平均值为零时,变异系数无意义。

本文试建议的均同指数与概率分布的散度特征单因素相关,除能够确切反映概率分布的离散程度外,还适用于统一评价或比较不同子样的离散条件。

6件岩土样本要求只是相应推荐公式的适用界限,不能作为控制样本数量的执行标准。

相对容量控制原则是结合岩土体的空间分布条件,按满足相应概括性的抽取率标准科学合理限定岩土样本数量的方法。建议相应规范或标准进一步明确限定岩土样本采集技术工作勘探点位的水平间隔和沿深度方向的采样间距,实际工作可根据实际采取样本数量与相应岩土层中的勘探累计总进尺,综合考评相关岩土工程勘察的工作质量。具体相对容量标准,尚待通过进一步的经验积累逐步完善和制定。

建议规范将标准值计算近似公式适用范围明确限定为实际参与统计推断的样本数量 $\geq 5$ 件,样本数量为4件时,应结合样本的均同性等条件慎重采用。更小样本容量条件下标准值的推断方式,须由岩土工程师根据经验结合样本的概括性和岩土母体的离散程度及其他工程背景条件综合分析确定。

文献[4]建议的修正计算公式,实质是依据岩土参数平均值偶然性置信估计确定的标准值。建议切实结合岩土母体所固有的分布特征,按90%设置相应的置信水平统计推断岩土参数标准值。

岩土样本离散程度除与成因环境及岩土参数类别有关外,一定程度上还受岩土母体分布范围分划条件的影响。因此岩土参数标准值的统计分析,应依据岩土样本的离散等级等条件划分的统计单元进行,重要性的建设工程应限定在建设工程场地影响范围内划分统计分区。同时标准值修正的侧限关系,应结合可靠性分析意图,按最低风险控制或最不利可能区别对待。

(下转第122页)

表2 边坡稳定性计算表

条块编号	条块自重 ( $\text{kN} \cdot \text{m}^{-1}$ )	条块底面 倾角 $\theta_i/(\circ)$	内摩擦角 $\varphi_i/(\circ)$	粘聚力 $c_i/\text{kPa}$	条块滑动面 长度 $l_i/\text{m}$	安全系数 $\gamma_t$	传递系数 $\psi$	剩余下滑力 ( $\text{kN} \cdot \text{m}^{-1}$ )	稳定性 系数 $K_s$
1	17.9	62.5	17	18	2.18	1.25	0.994	0.0	
2	185.1	61.9	28	18	415.88	1.25	0.991	69.9	
3	310.3	61.2	35	20	3.93	1.25	0.991	225.9	
4	480.4	60.5	35	20	4.06	1.25	0.996	500.8	
5	563.6	59.9	20	20	3.58	1.25	1.027	949.5	
6	785.8	63	28	90	4.41	1.25	1.000	1238.2	0.65
7	966.1	63	28	90	4.41	1.25	1.000	1684.0	
8	1146.3	63	28	90	4.41	1.25	1.000	2287.1	
9	1326.8	63	28	90	4.41	1.25	1.000	3047.7	
10	1498.2	63	28	90	4.41	1.25	1.000	3957.7	
11	1428.8	63	28	90	6.39	1.25	1.000	4629.0	

### 3 结论

1) 土-岩混合边坡稳定分析的特点是土层与岩层即相互独立又相互关联,不宜将两者单独孤立分析。

2) 可能破坏形式存在多种,土体沿内部产生局部圆弧滑动、土体沿土岩分界面滑移,当岩质部分不稳定时,则边坡沿软弱结构面带动土体整体滑移。

3) 分析某一特定边坡时,应考虑土体局部破坏和岩体的结构面发育组合情况,判断是否有条件发生破坏。

4) 提出了土体最危险潜在滑面与下部岩层中结构面组合,采用传递系数法进行整体稳定性分析的思路和方法。

### 参 考 文 献

- [1] 中华人民共和国建设部. GB 50330—2002 建筑边坡工程技术规范[S]. 北京: 中国建筑工业出版社, 2002.
- [2] 朱强, 刘伟. 土岩混合高边坡锚索(杆)加固前后稳定性分析[J]. 勘察科学技术, 2007(3): 13-16.
- [3] 冯君, 周德培. 顺层岩质边坡顺层滑动岩体范围分析[J]. 山地学报, 2007, 25(3): 376-380.
- [4] 周中, 傅鹤林. 土石混合体边坡人工降雨模拟试验研究[J]. 岩土力学, 2007, 28(7): 1391-1396.
- [5] 钟正雄, 黄志洲, 储洪宝. 岩石边坡稳定性分析研究方法及其讨论[J]. 河海大学学报, 2001, 29(增刊): 69-71.

收稿日期: 2008-01-31

(上接第112页)

### 5 结 语

岩土母体的特殊性,是概率统计的实际应用需要经历持续改进过程的主要原因。本文所阐述观点,源于笔者多年从事岩土工程专业技术工作的感悟,限于自身基础理论水平和专业技术能力,文中难免存在偏颇或谬误,敬请各位同行、学长给予批评和指正。

致谢: 谨此此文诚挚感谢高大钊老师的赠书指教及同济大学第十一期岩土工程短训班期间的教示。

### 参 考 文 献

- [1] 学教材[M]. 北京: 人民教育出版社, 1980: 149-155.
- [2] 《数学手册》编写组. 数学手册[M]. 北京: 人民教育出版社, 1979: 809-810.
- [3] 高大钊. 土力学可靠性原理[M]. 北京: 中国建筑工业出版社, 1989: 23.
- [4] GB 50021—2001 岩土工程勘察规范[S].
- [5] 常士骠, 张苏民主编. 工程地质手册(第三版)[M]. 北京: 中国建筑工业出版社, 1992.
- [6] 林宗元主编. 岩土工程勘察设计手册[M]. 沈阳: 辽宁科学技术出版社, 1996.
- [7] JGJ 72—2004 高层建筑岩土工程勘察规程[S].

收稿日期: 2008-02-13