

基于有限比较法研究万州区滑带土干重度的概率分布

刘莲花, 罗文强

(中国地质大学数学与物理学院, 湖北 武汉 430074)

摘 要: 对万州区滑坡滑带土干重度的 86个测试数据进行拟合优度检验, 发现多种传统的检验方法都既可以接受正态分布也可以接受对数正态分布。在传统检验方法失效的情况下, 引入了有限比较法, 确定了干重度的最优概率分布为正态分布, 并用统计模拟的方法分析了有限比较法选中真实分布的概率。

关键词: 干重度; 概率分布; 拟合优度检验; 有限比较法; 统计模拟

中图分类号: P642. 11, TU 41

文献标识码: A

在可靠性分析中, 参数的概率分布类型是破坏概率和可靠指标计算的基础, 直接影响着可靠指标的计算结果。例如抗剪强度参数内聚力和内摩擦角取不同概率分布类型的随机变量, 当他们的期望值和方差相同时, 当量正态化后计算的可靠度指标值相差达 20% ~ 30%, 对概率分布类型较为敏感^[1], 因此, 对参数概率分布的研究具有重要的意义。目前的研究成果主要是利用某一种拟合优度检验, 确定参数的概率分布为正态分布、对数正态分布等^[2-3]。但是这种传统的检验方法结果容易出现失效的情况, 例如: 对于同一样本在某些显著性水平下, 无论用什么检验方法, 多种性质不同的原分布假设均不被拒绝, 这种情况下, 参数的分布类型就不能确定。

本文结合三峡库区万州段三期地质灾害规划勘察的 27个滑坡实例测试参数, 采用有限比较法, 给出了传统检验方法失效情况下, 参数分布类型的确定方法和选择标准, 并用统计模拟的方法分析了有限比较法的合理性。

1 实例研究

万州区位于三峡库区中段, 大型滑坡群有 7 个^[3]。大量滑坡的勘察和治理为滑坡滑带土参数的概率分布研究提供了丰富的资料, 由于样本量较大, 为研究该地区滑带土参数的概率分布提供了可能。选者资料^[4-5]中滑带物质组成相近的 27处滑坡, 统计了 86个滑带土干重度的测试值(表 1), 要求根据测试值确定干重度的概率分布。

根据实践经验和前人研究成果, 干重度可能服从正态分布也可能服从对数正态分布, 因此我们可以提出如下的两种原假设(显著性水平 $\alpha = 0.05$):

H 01: 干重度服从正态分布

H 02: 干重度服从对数正态分布

对于上述两种原假设, 我们可采用传统的拟合优度检验方法进行检验。常用的传统的拟合优度检验^[6]有卡方检验, QQ图检验以及 K-S, C-M, A-D 等多种“EDF”检验法, 还有专门用于正态性检验的 W 检验法和 D 检验法, 本文采用开方检验法、C

收稿日期 (Received date): 2008- 04- 15; 改回日期 (Accepted): 2009- 03- 15.

基金项目 (Foundation item): 国家自然科学基金资助项目 (40372120, 40672189); 武汉市研究创新计划项目 (20066002059). [Supported by the National Natural Science Foundation of China (No. 40372120, 40672189) and Project of Research Innovation of Wuhan (No. 20066002059).]

作者简介 (Biography): 刘莲花 (1983-), 女 (汉族), 湖北广水人, 硕士, 现为中国地质大学 (武汉) 应用数学专业研究生, 主要从事概率统计及应用方面的研究工作。 [Liu Lianhua (1983-), female, born in Guangshui, Hubei province. Graduate student of China University of Geosciences (Wuhan), Major in Probability Statistics and its Application.] E-mail: cugbtus@yahoo.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

表 1 万州区滑带土干重度测试值

Table 1 Test data of dry unitweight of landslide zone soil in Wanzhou area

(kN/m³)

| | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 15. 090 | 15. 480 | 16. 370 | 16. 760 | 15. 970 | 15. 780 | 13. 622 | 13. 622 | 14. 308 | 14. 602 | 14. 406 |
| 14. 602 | 15. 390 | 16. 860 | 16. 170 | 15. 480 | 15. 580 | 15. 190 | 15. 580 | 14. 900 | 16. 760 | 15. 780 |
| 15. 580 | 15. 580 | 15. 680 | 16. 660 | 16. 070 | 14. 990 | 15. 880 | 14. 900 | 16. 100 | 16. 000 | 17. 200 |
| 16. 100 | 16. 700 | 16. 600 | 16. 000 | 14. 500 | 15. 400 | 14. 200 | 14. 900 | 14. 500 | 15. 000 | 15. 400 |
| 14. 300 | 15. 700 | 14. 900 | 15. 400 | 15. 300 | 15. 800 | 16. 700 | 16. 600 | 16. 300 | 16. 900 | 16. 200 |
| 16. 800 | 16. 800 | 17. 000 | 16. 400 | 17. 400 | 16. 000 | 16. 000 | 16. 600 | 15. 200 | 16. 100 | 17. 500 |
| 16. 500 | 16. 200 | 17. 000 | 16. 300 | 16. 100 | 15. 600 | 15. 800 | 16. 000 | 15. 700 | 15. 600 | 15. 400 |
| 15. 700 | 15. 300 | 15. 300 | 15. 500 | 15. 500 | 15. 600 | 15. 580 | 15. 700 | 15. 600 | | |

- M 检验法和 A - D 检验法进行干重度概率分布的检验, 检验的结果如表 2 所示。

结果分析: 从表 2 可看出三种传统的拟合优度检验方法在两种不同的原假设分布下的统计量值均小于所对应临界值, 即三种检验方法都既可以接受正态分布的原假设又可以接受对数正态分布的原假设, 显然传统的检验方法无法给出干重度概率分布的确切结论。

这是因为用传统的拟合优度检验方法进行检验时可能犯两类错误: “弃真”和“取伪”。“弃真”的概率一般都是固定的且很小, 而“取伪”的概率却是不确定的, 并且当显著性水平给定时, “取伪”的概率与样本容量的大小成反比, 当检验的样本容量较小时就很容易犯“取伪”的错误, 从而接受多种性质不同的原假设分布, 即实例中出现的情况, 而要减小“取伪”的概率, 只有增大样本容量, 且当样本容量趋于无穷时, “取伪”的概率趋于零, 这时可唯一地确定样本的概率分布, 但是这在实践中是行不通的。这时我们应怎样来确定干重度的概率分布呢? 本文引入了如下的有限比较法。

2 有限比较法

方法介绍: 当用传统的拟合优度检验方法进行检验时, 可能会出现同时接受几种不同原假设分布的情况, 但是根据样本与母体的关系, 应该是符合母体的原假设分布拟合度最优。因此, 只需对有限几种可能的分布用传统的方法进行检验, 计算各自统计量的值, 然后选择统计量最小者对应的分布为母体的真实概率分布。这种以传统拟合优度检验方法为基础, 最终唯一确定样本概率分布的方法称为有

限比较法^[7]。

方法的模拟分析: 现有的研究方法多集中于有限比较法的应用, 而很少有讨论该方法的合理性, 即此方法是否一定能选中真实的概率分布, 本文采用统计模拟的方法进行分析。模拟中分别以实践中应用较广的正态分布和对数正态分布之一为真实分布进行抽样, 然后对这些已知其母体分布的样本用有限比较法进行检验, 看看能否选中真实的母体分布。

每次模拟中抽样总次数 $m = 2000$ 选中真实母体分布的频数为 m_0 , 由大数定律知, 当 m 足够大时, 可近似认为 m_0/m 就是选中真实分布的概率, 模拟结果如表 3~ 表 5 所示。

从统计模拟结果 (见表 3~ 5) 可以看出: 使用有限比较法确定最优分布函数时, 选中真实分布的概率并非为 1, 并非如文献^[7]所说一定可以选中真实分布。因为母体和样本存在差异, 有限比较法选中的“母体分布”只是和样本偏差最小的概率分布而可能为非真实的母体分布, 其结果必然会有偏差, 但是真实分布的中选概率比较大, 说明此方法还是可取的。

表 2 万州区滑带土干重度在两种不同原假设下的检验结果

Table 2 Goodness of fit test results of dry unit weight under two different null hypotheses

| 两种原假设 | 检验方法 | | |
|--------------------|-----------------------|---------------------------|---------------------------|
| | 开方检验统计量值 (区间数为 15) | C - M 检验法 统计量 W^2 的值 | A - D 检验法 统计量 A^2 的值 |
| 正态分布 | 15. 244 1 | 0. 062 4 | 0. 388 3 |
| 对数正态分布 | 19. 659 0 | 0. 075 6 | 0. 487 9 |
| 临界值 ^[6] | 21. 026 1 | 0. 126 0 | 0. 752 0 |

表 3 统计量为 W^2 时不同母体参数的模拟结果 (样本量 $n = 100$)

Table 3 Simulation results of different population parameters using statistics W^2 (sample size $n = 100$)

| 母体分布 | 母体分布 具体概型 | 真实分布 频数 m_0 | 真实分布 频率 m_0/m | 母体分布 具体概型 | 真实分布 频数 m_0 | 真实分布 频率 m_0/m |
|--------|-----------------|------------------|--------------------|------------------|------------------|--------------------|
| 正态分布 | $N(15, 3.0^2)$ | 167 0 | 0.835 | $N(15, 2.0^2)$ | 152 2 | 0.761 |
| | $N(15, 2.0^2)$ | 152 2 | 0.761 | $N(14, 2.0^2)$ | 152 4 | 0.762 |
| | $N(15, 1.5^2)$ | 137 8 | 0.674 | $N(13, 2.0^2)$ | 150 1 | 0.751 |
| | $N(15, 1.0^2)$ | 124 9 | 0.625 | $N(12, 2.0^2)$ | 161 3 | 0.807 |
| | $N(15, 0.9^2)$ | 123 1 | 0.616 | $N(11, 2.0^2)$ | 165 0 | 0.825 |
| | $N(15, 0.8^2)$ | 124 2 | 0.621 | $N(10, 2.0^2)$ | 168 3 | 0.842 |
| 对数正态分布 | $LN(4, 0.5^2)$ | 198 0 | 0.990 | $LN(7, 0.2^2)$ | 163 6 | 0.818 |
| | $LN(4, 0.4^2)$ | 193 4 | 0.967 | $LN(6, 0.2^2)$ | 163 9 | 0.820 |
| | $LN(4, 0.3^2)$ | 182 9 | 0.915 | $LN(5, 0.2^2)$ | 164 8 | 0.824 |
| | $LN(4, 0.1^2)$ | 135 8 | 0.679 | $LN(4, 0.2^2)$ | 166 6 | 0.833 |
| | $LN(4, 0.09^2)$ | 132 7 | 0.664 | $LN(3.5, 0.2^2)$ | 168 3 | 0.842 |
| | $LN(4, 0.07^2)$ | 128 7 | 0.644 | $LN(3, 0.2^2)$ | 168 1 | 0.841 |

表 4 统计量为 A^2 时不同母体参数的模拟结果 (样本量 $n = 100$)

Table 4 Simulation results of different population parameters using statistics A^2 (sample size $n = 100$)

| 母体分布 | 母体分布 具体概型 | 真实分布 频数 m_0 | 真实分布 频率 m_0/m | 母体分布 具体概型 | 真实分布 频数 m_0 | 真实分布 频率 m_0/m |
|--------|-----------------|------------------|--------------------|------------------|------------------|--------------------|
| 正态分布 | $N(15, 3.0^2)$ | 745 | 0.873 | $N(15, 2.0^2)$ | 149 1 | 0.746 |
| | $N(15, 2.0^2)$ | 149 1 | 0.746 | $N(14, 2.0^2)$ | 153 9 | 0.770 |
| | $N(15, 1.5^2)$ | 143 2 | 0.716 | $N(13, 2.0^2)$ | 159 3 | 0.797 |
| | $N(15, 1.0^2)$ | 125 6 | 0.628 | $N(12, 2.0^2)$ | 161 5 | 0.808 |
| | $N(15, 0.9^2)$ | 122 0 | 0.610 | $N(11, 2.0^2)$ | 164 7 | 0.824 |
| | $N(15, 0.8^2)$ | 122 3 | 0.612 | $N(10, 2.0^2)$ | 172 2 | 0.861 |
| 对数正态分布 | $LN(4, 0.5^2)$ | 198 7 | 0.994 | $LN(7, 0.2^2)$ | 171 8 | 0.859 |
| | $LN(4, 0.4^2)$ | 196 0 | 0.980 | $LN(6, 0.2^2)$ | 171 2 | 0.856 |
| | $LN(4, 0.3^2)$ | 187 4 | 0.937 | $LN(5, 0.2^2)$ | 171 3 | 0.857 |
| | $LN(4, 0.1^2)$ | 141 3 | 0.707 | $LN(4, 0.2^2)$ | 171 1 | 0.856 |
| | $LN(4, 0.09^2)$ | 134 6 | 0.673 | $LN(3.5, 0.2^2)$ | 172 7 | 0.863 |
| | $LN(4, 0.07^2)$ | 131 3 | 0.657 | $LN(3, 0.2^2)$ | 168 0 | 0.840 |

表 5 统计量为 W^2 时不同样本量的模拟结果

(母体为 $N(15, 3.02)$)

Table 5 Simulation results of different sample size using statistics W^2

(population is $N(15, 3.02)$)

| 样本量 n | 真实分布 频数 m_0 | 真实分布 频率 m_0/m | 样本量 n | 真实分布 频数 m_0 | 真实分布 频率 m_0/m |
|------------|------------------|--------------------|------------|------------------|--------------------|
| 50 | 150 4 | 0.752 | 100 | 169 3 | 0.847 |
| 60 | 155 1 | 0.776 | 120 | 174 8 | 0.874 |
| 70 | 162 5 | 0.813 | 140 | 175 6 | 0.878 |
| 80 | 166 1 | 0.831 | 160 | 179 7 | 0.899 |
| 90 | 167 6 | 0.838 | 180 | 180 5 | 0.903 |

表 3和表 4表明: 真实分布的中选频率随母体参数的变化而改变, 无论母体为正态分布还是对数正态分布, 当均值固定时, 随着标准差的减小, 真实分布中选频率呈减小趋势; 当标准差固定时, 随着均值的减小, 真实分布中选频率呈增大趋势。表 3、表 4对比可知, 不同的检验统计量也会影响真实分布的中选频率。表 5表明: 当样本容量逐步增大时, 真实分布的中选频率也随之增大。

实例运用: 运用上述介绍的有限比较法再次对实例中的干重度进行拟合优度检验, 从表 2中的计算结果可看出, 三种不同的检验方法中, 正态分布所对应的统计量的值均小于对数正态分布所对应的统

计量的值, 所以三种方法结论一致, 均认为干重度服从正态分布。

3 结论

对万州区滑带土干重度进行拟合优度检验, 在传统的方法失效的情况下引入了有限比较法, 并通过统计模拟分析了有限比较法的合理性, 模拟结果表明: 有限比较法选中真实分布的概率并非为 1, 其概率大小与统计量的形式、母体的分布类型及参数、子样容量大小等因素有关, 但是选中真实分布的概率还是比较大的。最后通过有限比较法确定了干重度的最优概率分布为正态分布。

参考文献 (References)

- [1] Yan Chunfeng, Liu Dongyan, Zhang Jianhui *et al*. The susceptibility analysis of reliability for the probability distribution types of parameters in strength criterion [J]. *Chinese Journal of Rock Mechanics and Engineering*, 1999, 18(11): 36~39 [严春风, 刘东燕, 张建辉, 等. 岩土工程可靠度关于强度参数分布函数模型的敏感性分析 [J]. 岩石力学与工程学报, 1999, 18(11): 36~39]
- [2] Ni Wankui, Han Qibing. Statistical analysis of physical and mechanical indexes of the typical loess [J]. *Journal of Engineering Geology*, 2001, 9(1): 62~67 [倪万魁, 韩启龙. 黄土土性参数的统计

- 分析 [J]. 工程地质学报, 2001, 9(1): 62~67]
- [3] Luo Chong, Yin Kunlong, Chen Lixia *et al*. Probability distribution fitting and optimization of shear strength parameters in sliding zone along horizontal-stratum landslides in Wanzhou city [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2005, 24(9): 1588~1593 [罗冲, 殷坤龙, 陈丽霞, 等. 万州区滑坡滑带土抗剪强度参数概率分布拟合及其优化 [J]. 岩石力学与工程学报, 2005, 24(9): 1588~1593]
- [4] Hubei Geological Environmental Centre. Database of sliding zone in Chongqing [R]. Wuhan: Hubei Geological Environmental Centre, 2005 [湖北省地质环境总站. 重庆滑带数据库 [R]. 武汉: 湖北省地质环境总站, 2005]
- [5] Chongqing Yangtze River Institute of Geological Engineering Exploration. Geology exploration report of controlling landslides for the third plan in Wanzhou city of Three Gorges reservoir zone [R]. Chongqing: Chongqing Yangtze River Institute of Geological Engineering, 2003 [重庆长江工程勘察院. 三峡库区三期地质灾害防治万州区崩塌滑坡规划调(勘)查报告 [R]. 重庆: 重庆长江工程勘察院, 2003]
- [6] Yang Zhenhai. Goodness of Fit Test [M]. Anhui: Anhui Education Press, 1987 [杨振海. 拟合优度检验 [M]. 安徽: 安徽教育出版社, 1987]
- [7] Zhang Boting. Fitting and optimization through finite contrast method [J]. *Chinese Journal of Geotechnical Engineering*, 1991, 13(6): 84~91 [张博庭. 用有限比较法进行拟合优度检验 [J]. 岩土工程学报, 1991, 13(6): 84~91]

The Study of Probability Distribution of Dry Unit Weight of Landslide Zone Soil in Wanzhou Area on Finite Contrast Method

LIU Lianhua, LUO Wenqiang

(School of Mathematics and Physics, China University of Geosciences, Wuhan 430074, Hubei, China)

Abstract To determine the probability distribution of dry unit weight of landslide zone soil in Wanzhou area, this paper firstly apply several traditional goodness of fit methods and find that both normal distribution and logarithmic normal distribution are acceptable to a variety of traditional goodness of fit methods. In order to determine the optimal probability distribution of dry unit weight, finite contrast method is proposed. Statistical simulation was used to analyze the probability of choosing the real distribution while using finite contrast method. The simulation results show that the probability of choosing the real distribution is not equal to one, it is affected by several factors, such as the statistics pattern, the population distribution style, the population parameters, the sample size, etc. Though we may choose the unreal distribution while adopting finite contrast method, the probability of choosing the real one is much bigger, therefore finite contrast method is quite effective. In the end, finite contrast method is applied to determine the probability distribution of dry unit weight and derive a conclusion that the probability distribution of dry unit weight is normality.

Key words dry unit weight; probability distribution; goodness-of-fit test; finite contrast method; statistical simulation