华南农业大学学报 15 (3) 1994: 141～144
J.South China Agr.Univ.

用计算机随机模拟方法研究水田土壤圆锥指数

陈联诚 区颖刚
(华南农业大学工程技术学院，广州，510642)

摘要 本文将计算机随机模拟法应用到估计水田土壤圆锥指数所需抽样容量的研究中，是一次新的尝试，结果表明，对服从威布尔分布的水田土壤圆锥指数、正态分布的抽样容量公式仍然适用。在具有统计性质的农业工程领域的问题中，计算机随机模拟方法作为实际试验的补充，能大大减少费时费钱的试验次数，有着广泛的应用前景。

关键词 随机模拟法；计算机；圆锥指数；正态分布；威布尔分布
中图分类号 S152.9

随机模拟(Random Simulation)方法又叫蒙特卡罗(Monte－Carlo)方法，是一种以概率统计为基础，以随机抽样为主要手段的模拟方法，由于必须进行大量的随机抽样和计算才能得到可靠的结果，这种方法在近年随着计算机的普遍应用，才在数学物理，工程技术，生产管理各种不同问题中得到日益广泛的应用(徐钟济，1985; 张建中，1974; 裴鹿成，1982; 徐光辉等，1974)。而使用计算机实现模拟试验，则成了这种方法的基本特征。

本文应用随机模拟方法研究对水田土壤圆锥指数进行估计的问题，圆锥指数是土壤机械强度的综合指标值，在旱地土壤强度估计中的效果已被广泛接受(吴起亚等，1985)。在水田中的适用性还需要进一步探讨(苏显添，1985)。研究表明，水田土壤圆锥指数的分布服从威布尔分布(Weibull Distribution)，而不是正态分布(Normal Distribution)*。在这种情况下，估计一块水田的平均圆锥指数值，最少需要测定多少次，是一个需要进一步研究的问题，本文试图通过应用随机模拟方法对这个问题作一些探讨。

1 水田土壤圆锥指数抽样容量的理论分析

1.1 抽样理论

不论总体 X 为何种分布，用抽样平均值 $\bar{x}=\sum x/n$ 来估计总体平均值 μ，总是适用的。问题只在于对于不同的总体分布，要达到同样精度所要求的抽样容量 n 可能不同。

对于服从正态分布 $N=(\mu,\ \sigma^2)$ 的随机变量 x，其抽样平均值 $\bar{x}$ 也服从正态分布，且在95% 的置信度下，有

$$|\bar{x}-\mu|\leqslant 1.96\sigma/\sqrt{n}=\delta \quad\cdots\cdots(1)$$

1993－11－15 收稿

*Ou Y, Shao. Y. 1988. Statistical analysis of cone index readings in paddy field soils. In: B. D.Witney (eds). Tillage and traffic in crop production (Processings of 11th international conference of ISTRO, Edingburgh,Scotland), 97～102

式中:

$\delta = 1.96\sigma/\sqrt{n}$ 是测定精度,是根据专业要求预先确定的

取

$$\delta = d \cdot \mu \quad (2)$$

式中 d 为精度系数。

则:

$$\begin{aligned} n &= (1.96\sigma)^2/\delta^2 \\ &= (1.96\sigma)^2/(d^2 \cdot \mu^2) \\ &= (1.96 \cdot cv)^2/d^2 \quad (3) \end{aligned}$$

式中

$cv = \sigma/\mu$ 是变异系数,反映了数据的相对变化情况,在土壤机器系统力学的测试中,精度系数 d 可取为 0.2。

则

$$\begin{aligned} n &= (1.96 \cdot cv)^2/d^2 \\ &= 96 \cdot cv^2 \quad (4) \end{aligned}$$

1.2 在威布尔总体中随机抽样的方法

对于服从威布尔分布的土壤圆锥指数,未见有抽样容量公式可遵循,根据中心极限定理,任意分布的随机变量 X,只要抽样容量 n 充分大,其算术平均值 $\bar{x}$ 的分布就是近似正态分布,可按正态分布的理论处理。一般认为取 $n=30$ 即算充分大。因此,若按(4)式算得 $n>30$,则可按此 n 值抽样来计算圆锥指数。

对于 $n<30$ 的情况,上式是否合适,下面用随机模拟方法进一步探讨。

应用随机模拟方法前,首先要确定随机抽样的方法。

由随机模拟理论可知:连续分布的随机变量 η,若其分布函数为 $F(X)$, 则随机变量

$$R = F(\eta) \quad (5)$$

是 [0,1] 上均匀分布的随机变量,若分布函数 $F(x)$ 的反函数能求得,

$$\eta = F^{-1}(R) \quad (6)$$

则对在[0,1] 上均匀分布的 R 随机抽样,就可得到 η(徐钟济,1985)。

根据这个理论,试推出威布尔分布的直接抽样法:

威布尔分布的分布函数为

$$\begin{aligned} F &= F(x) \\ &= 1 - \exp[-(x-r)^m/t_0] \\ (x-r)^m &= t_0 \ln[1/(1-F)] \\ x &= [t_0 \ln[1/(1-F)]^{1/m} + r \quad (7) \end{aligned}$$

由计算机的随机函数,可产生服从[0,1] 均匀分布的随机的 F 值。由(7)式即可得到服从威布尔分布的模拟抽样值 x。

2 水田土壤圆锥指数抽样的计算机随机模拟

为研究水田土壤圆锥指数的概率分布,作者曾在水田土槽作过大量抽样、证实了它们确是服从威布尔分布,结果见表 1。

对这些分布进行计算机随机模拟抽样,先按(4)式计算所需抽样数,所得结果见表 2。

表 1　圆锥指数威布尔分布参数及 χ^2 检验值

组号	土层 /cm	样本数	威布尔参数 (m)	(t_0)	(r)	圆锥指数均值	方差 (σ)	χ^2 值	自由度
1	0～10	200	3.49	129.0	0.00	3.62	1.15	10.24	7
2	11～20	200	1.95	15.5	2.35	5.97	1.93	13.80	7
3	21～30	200	3.15	50.0	2.00	5.10	1.08	6.77	7
4	0～15	200	2.15	15.2	1.35	4.49	1.54	9.83	7
5	16～30	200	2.33	17.0	2.38	5.37	1.36	10.71	7
6	0～30	200	2.40	14.0	2.30	4.96	1.18	8.95	7

算得的 n 值都小于 30，若这此抽样容量 n 适合的话，则按这些 n 值对每一个威布尔分布总体进行模拟抽样时，每抽样 100 组(每组 n 个样本)，应有大约 95 组的平均值 $\bar{x}$ 落在 $(\mu-\delta, \mu+\delta)$ 范围内。

对表 1 的 6 个威布尔分布总体进行计算机模拟抽样，所用的直接抽样方法已如上述。每个分布抽 100 组，每组 n 个样本，结果及合格率见表 2。

表 2　计算抽样容量及计算机随机模拟试验结果

组号	计算抽样容量 (n)	精度系数 (d)	土层 /cm	变异系数 (cv)	分布均值 (μ)	合格范围	合格率 / %
1	10	0.2	0～10	0.32	3.62	2.896～4.345	96
2	10	0.2	11～20	0.32	5.97	4.773～7.159	96
3	4	0.2	21～30	0.21	5.10	4.079～6.118	93
4	11	0.2	0～15	0.34	4.49	3.592～5.388	93
5	6	0.2	16～30	0.25	5.37	4.295～6.443	93
6	5	0.2	0～30	0.24	4.96	3.970～5.955	89

3　结果讨论

由表 2 可以看出，按(4)式计算的 n 值抽样，合格率基本上是符合要求的。这说明对于威布尔分布来说，抽样公式(4)仍然是适用的。影响抽样次数,主要是数据的变异系数，而分布的形式是次要的。

为了进一步验证上述计算机模拟结果的正确性，直接用水田土槽试验中在 1.6 m ×10 m 面积上大密度抽样得到的6 组(每组 200 个)圆锥指数值作为 6 个总体，分别取 n＝1、2、3、4、5、10，在计算机上对每组 200 个数进行随机抽样，每个总体每个 n 值抽 100 组(每组 n 个样本)。结果如表 3。结果表明，(4)式计算的 n 值仍是适当的。这也说明，计算机随机模拟的结果是可靠的。如果不采用计算机模拟法，而是直接到田间抽样，要完成上述两次抽样检验，共需测定 6 万多次，其工作量是不可想象的。

4　结论

对服从威布尔分布的水田圆锥指数的计算机随机模拟抽样表明，影响抽样次数，主要是数据的变异系数，而分布形式是次要的。正态分布总体的抽样公式在这种情况下仍然适用。计算机随机模拟方法是一种对具有统计性质的物理现象进行深入研究的有力工具。它

表 3 计算机随机抽样检验结果合格率

抽样容量	土		层			
n	0 ~ 10 cm	11 ~ 20 cm	21 ~ 30 cm	0 ~ 15 cm	16 ~ 30 cm	0 ~ 30 cm
1	0.47	0.39	0.57	0.46	0.67	0.63
2	0.56	0.54	0.88	0.66	0.76	0.79
3	0.71	0.70	0.87	0.67	0.83	0.89
4	0.81	0.72	0.94	0.79	0.88	0.92
5	0.82	0.80	0.97	0.88	0.93	0.96
6					0.96	0.99
10	0.99	0.96	1.00	0.96	0.99	1.00

作为实际试验的补充，能大大减少费时费钱的试验次数。本研究如果进行实际试验，则要在田间抽样上万次，才能得到同样的结论。

参 考 文 献

徐钟济．1985. 蒙特卡罗方法．上海：科学技术出版社，1 ~ 10. 13, 25 ~ 28, 101

张建中．1974. 蒙特卡洛方法：Ⅰ．数学的实践与认识，(2)：43 ~ 56

裴鹿成．1982. 伴随蒙特卡洛方法在屏蔽计算中的应用．数值计算与计算机应用，(1)：24

徐光辉，董泽清．1974. 矿山采掘过程的数字模拟．数学的实践与认识，(4)：26 ~ 37

吴起亚，高行方．1985. 拖拉机与农业机械牵引力学．北京：中国农业机械出版社，15 ~ 16

苏显添．1985. 试论圆锥指数的适用性．农业机械学报，(1)：12 ~ 23

THE APPLICATION OF COMPUTER RANDOM SIMULATION METHOD TO PADDY SOIL CONE－INDEX

Chen Liancheng   Ou Yinggang

(College of Polytechnic, South China Agr. Univ. Guangzhou, 510642)

Abstract

The computer random simulation method was applied, as a new attempt, to investigate the sampling size required in estimating the paddy soil cone－index in this paper. The results indicated that the sampling size formula for Normal Distribution was applicable to paddy soil cone－index which follows Weibull distribution. For problems with statistical characters, the computer random simulation method can help to reduce enormously the number of time and money－consuming experiments required, and be widely used in the area of agricultural engineering.

Key words random simulation method; computer; cone－index; normal distribution; Weibull distribution