

方法介绍

数据分布类型检验及其在土壤学中的应用

III. Vistelius 置信带法及其计算机程序

唐 涌 六

(中国科学院南京土壤研究所)

Vistelius 置信带法系 Vistelius 氏于 1960 年首先用于地质样品中微量元素浓度分布的正态性检验法[1]。嗣后,被用来检验土壤中微量元素的浓度分布类型[2]。由于该法的应用较晚,且计算过程极其冗繁,因此不为人们所熟知。然而,该法有着明显的优点,即判别精度高,它的辨别能力按置信水平计可分为 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 等多级。偏度、峰度法只有 0.02, 0.1 两级[3], Shapiro-Wilk W 检验法只有 0.01, 0.02, 0.05, 0.1, 0.5 五级[4]。后两种方法对置信水平大于 0.1 的样本的辨别能力明显不足。就样本大小而言, Vistelius 置信带法对大样本及小样本均能适用,这又比 W 检验法优越。虽然有计算复杂的缺点,但藉电子计算机的帮助,应用此法并无困难。

检验的方法是在观测值 X 对 X 的标准正态变量 U 的坐标平面上标绘点 $U(X)$, 接着计算出经验函数增长点 $\widetilde{U}_i$ 的标准偏差 $\sigma\widetilde{U}_i$ 。其计算公式为:

$$U = \frac{X_i - \bar{X}}{S} \dots\dots\dots (1)$$

$$\sigma\widetilde{U}_i = \frac{1}{g(U_i)} \sqrt{\frac{\phi(U_i) [1 - \phi(U_i)]}{n}} \dots\dots\dots (2)$$

式中, X_i 为样品观测值, $\bar{X}$ 为样本均值, S 为样本标准差, U 为 X_i 的标准正态变量。 n 为样品数。 $g(U_i)$ 为正态概率密度, $\phi(U_i)$ 为正态分布面积,两者均可根据 U 值从统计表中查得。接着,便可建立不同置信水平的置信带以检验样本在何种置信水平上遵从正态分布。例如规定置信系数 $P > 0.05$ 时接受假设,则如果所有的观测值全部落在置信系数 $P = 0.05$ 的置信带内,接受原假设。如有任一观测值落在该置信带之外,则拒绝原假设。由于一个对数正态样本的观测值的对数呈正态分布,因此,分别检验观测值及其对数值何者遵从正态分布,或何者具有较大的置信水平,即可判定该样本属何种分布。至于本检验方法的详细原理及计算步骤,可参阅文献[1]及[2]。本文仅介绍应用计算机解题,可按表 1 的步骤操作计算机即可。

本程序在运行时要求使用者输入样本名称,样品数,样品观测值,答复是否须要核对及须改正的样品

表 9 磷肥直接和间接施用对作物的增产效果

作物	每斤磷肥直接施用增产(斤)		每斤磷肥通过绿肥的肥效(斤)		
	幅 度	平 均	增 产	较直接施用多增产	肥效提高(%)
水 稻	1.8~3.6	2.7	5.4	2.7	100
小 麦	1.6~2.6	2.0	3	1	50
玉 米	1.1~2.5	1.3	3.6	2.3	177
棉花(皮棉)	0.4~1.6	1	2.8	1.3	130

三、结语

过磷酸钙在与泥肥共沤过程中,并未发现有明显的促进沤肥腐熟的作用,在酸性泥肥中有效磷被大量吸持。这种沤肥无论用于水田或旱地,都比将磷肥在栽

种时直接施用的效果差。磷肥与泥肥共沤,打乱了磷肥的合理布局,使得有限的磷肥不能按作物种类和土壤缺磷程度的需要施用,特别是影响了绿肥作物优先施用磷肥的优越性。综上所述,将过磷酸钙和泥肥共沤,从本试验所得的经济效益看,似乎没有这种必要性。

```

10:REM PROG.VISTE
LIUS
20:" ":CLEAR:
WAIT 0:INPUT "
ELEMENT="";A$
30:INPUT "NO.OF S
AMPLE="";N
35:LF 2:LPRINT A$
;" N="";N:LF 1
40:B=N-1
50:DIM X(B),W(B),
F(B),G(B),K(B)
,L(7),R(60)
60:L(0)=2.526:L(1)
)=2.326:L(2)=1
.96:L(3)=1.645
:L(4)=1.282:L(
5)=1.036:L(6)=
.842
65:L(7)=.672
70:FOR I=0TO B
80:INPUT "DATA=";
X(I)
90:NEXT I:CLS
100:INPUT "VERIFY?
(Y,N)";B$
105:IF (B$="Y")+ (B
$="N")<>1GOTO
100
110:IF B$="N"GOTO
100
120:FOR I=0TO B
130:LPRINT "DA";I+
1;"="";X(I)
140:NEXT I:LF 1
150:INPUT "CORR.NO
."";I:GOTO 170
160:GOTO 180
170:INPUT "DATA=";
X(I-1):GOTO 15
0
180:FOR I=0TO N-2
190:FOR J=0TO N-2-
I
200:IF X(J)<X(J+1
)GOTO 220
210:D=X(J):X(J)=X(
J+1):X(J+1)=D
220:NEXT J
230:NEXT I
270:P=1/J(2*PI):E
=0:R(15)=.5:U=
0:C=P:H=0.01
280:FOR I=1TO 15
290:FOR J=1TO 10
300:U=U+H:U=C+4*P*
EXP (-U*U/2):U

```

```

=U+H:C=P*EXP (
-U*U/2):E=E+(C
+U)/300
310:NEXT J
320:R(15-1)=0.5-E:
R(15+1)=0.5+E
330:NEXT I
340:FOR I=0TO B
350:E=(1+.5)/N:J=0
360:J=J+1:IF E<R(J
)GOTO 380
370:IF J<30GOTO 36
0
380:Z=R(J-1):U=0.2
*J-3.2:C=P*EXP
(-U*U/2):IF Z>
=EGOTO 430
390:U=U+H:U=C+4*P*
EXP (-U*U/2):U
=U+H:C=P*EXP (
-U*U/2):U=(C+U
)/300:Z=Z+U
400:IF Z<EGOTO 390
410:IF Z=EGOTO 430
420:U=U-(Z-E)/U*0.
02
430:W(I)=U
440:NEXT I
450:LPRINT "NORMAL
"
460:GOSUB 550
470:LF 1:LPRINT "L
N NORMAL"
480:FOR I=0TO B
490:X(I)=LN X(I)
500:NEXT I
510:GOSUB 550
520:END
550:M=0:S=0
560:FOR I=0TO B
570:M=M+X(I):S=S+X
(I)*X(I)
580:NEXT I
590:S=S/(S-M*M/N)/
B):M=M/N
600:LF 1:LPRINT "M
=";.00001*INT
(100000*M):
LPRINT "SD=";.
00001*INT (100
000*S)
610:FOR I=0TO B
620:U=(X(I)-M)/S:J
=INT (15+5*U):
E=.001
630:IF J<0GOTO 700
640:E=R(J):Q=.2*(J

```

```

-15)
650:IF Q=UGOTO 700
660:D=1+INT ((U-Q)
/.02):H=(U-Q)/
D:Z=H/2:T=H/6:
C=P*EXP (-Q*Q/
2)
670:FOR J=1TO D
680:Q=Q+Z:U=C+4*P*
EXP (-Q*Q/2):Q
=Q+Z:C=P*EXP (
-Q*Q/2):E=E+(C
+U)*T
690:NEXT J
700:F(I)=E:G(I)=.
EXP (-U*U/2)*P
:K(I)=(E*(1-E
)/N)/G(I)
710:NEXT I
720:I=0
730:J=0
740:H=L(I)*K(J)
750:IF W(J)<(X(J)-
M)/S-HGOTO 800
760:IF W(J)>(X(J)-
M)/S+HGOTO 800
770:J=J+1:IF J<N
GOTO 740
780:I=I+1:IF I<8
GOTO 730
790:LPRINT "P>0.5"
:GOTO 900
800:IF I=0LPRINT "
P<0.01":GOTO 9
00
810:IF I=1LPRINT "
0.02>P>0.01":
GOTO 900
820:IF I=2LPRINT "
0.05>P>0.02":
GOTO 900
830:IF I=3LPRINT "
0.1>P>0.05":
GOTO 900
840:IF I=4LPRINT "
0.2>P>0.1":
GOTO 900
850:IF I=5LPRINT "
0.3>P>0.2":
GOTO 900
860:IF I=6LPRINT "
0.4>P>0.3":
GOTO 900
870:LPRINT "0.5>P>
0.4"
900:RETURN

```

Sc N= 22

DA 1= 3.4
DA 2= 5.7
DA 3= 9.6
DA 4= 10.3
DA 5= 11.8
DA 6= 12.1
DA 7= 12.2
DA 8= 12.9
DA 9= 13
DA 10= 13
DA 11= 13.3
DA 12= 13.5
DA 13= 13.7
DA 14= 14.9
DA 15= 15
DA 16= 15.5
DA 17= 15.6
DA 18= 16.5
DA 19= 17.1
DA 20= 17.5
DA 21= 19.6
DA 22= 22.9

NORMAL

M= 13.59545
SD= 4.2073
0.2>P>0.1

LN NORMAL

M= 2.54657
SD= 0.40551
0.02>P>0.01

表 1 键 盘 操 作 步 骤

步 骤	输 入	显 示	说 明
1	DEF SPACE	ELEMENT =	问元素名称
2	SC ENTER	NO. OF SAMPLE =	问样品数
3	22 ENTER	DATA =	要求输入数据
4	3.4 ENTER	DATA =	要求继续输入数据
25	22.9 ENTER	VERIFY? (Y, N)	问是否要核对
26	N ENTER		印出计算结果, 终止
	Y ENTER	CORR. NO =	印出输入的数据供核对, 问须改正的样品号
27	ENTER*		印出计算结果, 终止

* 本例没有须要修正的数据, 如有须修改的数据, 其操作方法可参阅文献[3]。

号及其数值。计算完成后, 印出正态性检验的置信水平、算术平均值及标准差, 对数正态性检验的置信水平、观测值的对数值的平均值及标准差。使用者比较正态及对数正态检验的置信水平的大小, 即可判定样品属何种分布。为便于说明, 在程序的后面提供了一个例题, 这是南京地区22个土壤样品的浓度分布类型检验, 具体数据见语句900以后的印出部分。从印出部分可见, 正态检验的结果为 $0.2 > P > 0.1$ 而对数正态检验的结果为 $0.02 > P > 0.01$ 因此可断定样本服从正态分布而不属于对数正态分布。这个结论与文献[3]中用偏度、峰度法所得到的结论相同。

程序中的有关参数与符号, 说明如下:

X, 场, 存样品的观测值 X_i , N个向量。

W, 场, 存经验函数增长点, N个向量。

F, 场, 存正态分布面积, N个向量。

G, 场, 存正态概率密度, N个向量。

K, 场, 存经验函数增长点的标准偏差, N个向量。

L(7), 场, 存t分布的双侧分位数。

R(60), 场, 存正态积分函数。

A\$, 元素名称; B\$, 核对否。

N, 样品数; B = N - 1。

I, J, 循环参数。

C, D, E, M, Q, S, T, U, V, Z, 工作单元。

H, 步长。

P, 常数 $1/\sqrt{2\pi}$ 。

计算机的印出结果中, 有关符号说明如下:

N, 样品数。

Sc, 元素名称, 本例为钒。

DA, 数据, 即输入的样品观测值。

NORMAL, 正态性检验。

M, 均值。

SD, 标准差。

P, 置信水平。

LN NORMAL, 对数正态性检验。

参 考 文 献

- [1] Vistelius, A. B., J. Geol., 68: 1—22, 1960.
- [2] 唐涌六, 南京地区土壤中重金属浓度的概率分布。环境中若干元素的自然背景值及其研究方法, 9—15页, 科学出版社, 1982。
- [3] 唐涌六, 数据分布类型检验及其在土壤学中的应用 I. 偏度、峰度检验法及其计算机程序。土壤, 16卷2期, 66—69页, 1984。
- [4] 唐涌六, 数据分布类型检验及其在土壤学中的应用 I. Shapiro—Wilk W 检验法及其计算机程序。土壤, 16卷3期, 1984。