第 31 卷 第 6 期 西北农林科技大学学报(自然科学版) Vol. 31 No. 6
2003 年 12 月 Jour. of Northwest Sci-Tech Univ. of Agri. and For. (Nat. Sci. Ed.) Dec. 2003

毒力测定的统计分析方法研究*

袁志发[1], 卢恩双[1], 郭满才[1], 李 刚[2]

(1 西北农林科技大学 生命科学学院, 陕西 杨陵 712100; 2 西北农林科技大学 无公害农药研究服务中心, 陕西 杨陵 712100)

[摘 要] 在机率值法和新机率值法的基础上, 提出了基于Logistic 回归的 2 种方法, 并对各方法的异同进行了比较。结果表明, Logistic 回归法计算简单, 分析的准确度和机率值法一致。

[关键词] 毒力测定; 机率值; Logistic 回归; 致死中量

[中图分类号] S11+4 [文献标识码] A [文章编号] 1671-9387(2003)06-0181-03

毒力测定的试验设计是单因素试验, 考察因素是杀虫剂的不同剂量或浓度 $\lambda_1, \lambda_2, \cdots, \lambda_k$ (一般 $k=5$ 或 6), 供试的昆虫数量为 $n_1, n_2, \cdots, n_k$; 观察的指标为死亡率 $\hat{p}_i = \frac{\lambda_i}{n_i}$ (λ_i 为死亡的昆虫个数)。然而, 死亡率 p_i 不与剂量的增加成比例, 而是与剂量增加的比例成比例, 故 $p_i = \int^{\tau} f(\lambda)\mathrm{d}\lambda$ 中的密度函数 $f(\lambda)$ 为偏态, 累积死亡率和 λ 的关系是 1 条不对称的"S"型曲线。为了将分布变为对称分布, 把 $x_i = \lg\lambda_i$ 称为剂量 λ_i 的剂量值, 并假定 $X \sim N(m, \sigma^2)$, 其中 m 为杀死昆虫群体半数的剂量, 即 $m = \lg LD_{50}$, LD_{50} 称为致死中量; σ^2 为 X 的方差, σ^2 愈小, 反应愈激烈。毒力分析的目的是估计 m 和 σ^2, 从而估计 LD_{50} 及其方差。Gaddum 和 Bliss 提出了至今仍广泛使用的毒力测定机率值分析法[1,2], 卢恩双等[3]也对机率值分析方法提出了新的看法, 本研究提出用 Logistic 回归[4] (因变量为 0～1 分布) 分析法对毒力测定进行分析, 以克服新、旧机率值计算中的复杂性, 并对其准确度予以讨论。

1 Caddum 和 Bliss 的机率值分析法

假定剂量值 $X = \lg\lambda \sim N(m, \sigma^2)$, 则 $u = \frac{X-m}{\sigma} \sim N(0,1)$, 剂量 x_i 之下的理论死亡率为:

$$p_i = \Phi(u_i) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u_i} \mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u \tag{1}$$

当理论死亡数 np_i 和实际死亡数 r_i 使

$$\chi^2 = \sum_{i=1}^{k}\frac{(r_i - n_i p_i)^2}{n_i p_i} \sim \chi^2(k-2) \tag{2}$$

不显著时, 则有如下理论线性回归方程:

$$y_i = 5 + u_i = 5 + \frac{1}{\sigma}(x_i - m) = 5 - \frac{m}{\sigma} + \frac{1}{\sigma}x_i \tag{3}$$

$y \sim N(5, 1)$, Bliss 称 y 为机率值, y_i 可由已知的 $\hat{p}_i = \Phi(u_i)$ 反查出 u_i 而获得。

方程(3)可由 (x_i, y_i) $(i = 1, 2, \cdots, k)$ 进行回归分析而获得, $\hat{y} = a_0 + ax$, $\frac{1}{\sigma}$ 的估计值为 a, a_0 为 $5 - \frac{m}{\sigma}$ 的估计, 这样有:

$$\hat{\sigma} = \frac{1}{a}, \qquad \hat{m} = \frac{5 - a_0}{a} \tag{4}$$

由于在不同 x_i 时, 同一供试虫数 n 的死亡数不一样, 因而必须用死亡率 p_i 和存活率 $1 - p_i$ 的平均标准化剂量乘积来矫正:

$$w_i = \left(\frac{1}{p_i}\int_{-\infty}^{u_i}\frac{1}{\sqrt{2\pi}}u\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u\right) \cdot \left(\frac{1}{1-p_i}\int_{-\infty}^{u_i}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u\right) = \frac{1}{p_i(1-p_i)}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u_i^2}{2}}\right)^2 \tag{5}$$

然后用加权最小二乘法来估计 a_0 和 a:

---

* [收稿日期] 2003-06-13
[作者简介] 袁志发(1938-), 男, 陕西渭南人, 教授, 博士生导师, 主要从事农业应用数学研究。

$$\begin{cases} Q=\sum_{i=1}^{k} n_i w_i[(y_i-\overline{y})-b(x_i-\overline{x})]^2=\min \\ a=\dfrac{\sum n_i w_i(x_i-\overline{x})(y_i-\overline{y})}{\sum n_i w_i(x_i-\overline{x})^2}=\dfrac{l_{xy}}{l_{xx}} \\ a_0=\dfrac{\sum n_i w_i y_i}{\sum n_i w_i}-b\dfrac{\sum n_i w_i y_i}{\sum n_i w_i}=\overline{y}-b\overline{x} \end{cases} \tag{6}$$

其中, $\overline{x}=\dfrac{\sum n_i w_i x_i}{\sum n_i w_i}$, $\overline{y}=\dfrac{\sum n_i w_i y_i}{\sum n_i w_i}$,

$$l_{xx}=\sum n_i w_i x_i^2-\frac{(\sum n_i w_i x_i)^2}{\sum n_i w_i},$$

$$l_{xy}=\sum n_i w_i x_i y_i-\frac{(\sum n_i w_i x_i)(\sum n_i w_i y_i)}{\sum n_i w_i},$$

$$l_{yy}=\sum n_i w_i y_i^2-\frac{(\sum n_i w_i y_i)^2}{\sum n_i w_i}。$$

回归方程可用相关系数 $r=\dfrac{l_{xy}}{\sqrt{l_{xx}l_{yy}}}$进行显著性检验, 自由度为 $k-2$。回归方程显著时, 在 $x=\hat{m}$ 处的方差为:

$$S_{\hat{y}}^2=\left(\frac{1}{\sum n_i m_i}+\frac{(\hat{m}-\overline{x})^2}{l_{xx}}\right)\Big/a^2 \tag{7}$$

结合式(4)和式(7)有

$$\mathrm{LD}_{50}\pm S_{\mathrm{LD}_{50}}=10^{\hat{m}}\pm 10^{S_{\hat{y}}} \tag{8}$$

2 新机率值分析法

卢恩双等[3]认为, 用 $x=\hat{m}$ 处 $\hat{y}$ 的方差作为 $\hat{m}$ 的方差是不合适的, 应该用回归方程 $\hat{x}=b_0+by$ 来估计 m 及其方差, 其结果为:

$$\begin{cases} b=\dfrac{l_{xy}}{l_{yy}}, \qquad b_0=\overline{x}-b\overline{y} \\ \hat{m}=b_0+5b=\overline{x}+(5-\overline{y})b, \qquad \hat{\sigma}^2=b^2 \\ S_{\hat{m}}^2=b^2\left[\dfrac{1}{\sum n_i w_i}+\dfrac{(5-\overline{y})^2}{l_{yy}}\right] \\ \mathrm{LD}_{50}\pm S_{\mathrm{LD}_{50}}=10^{\hat{m}}\pm 10^{S_{\hat{m}}} \end{cases} \tag{9}$$

显然, 新、旧机率值法在显著性检验上是一致的。

3 Logistic 回归分析法

新、旧机率值法在分析计算时要由 $\hat{p}_i=\Phi(u_i)$ 反查出 u_i, 然后才能查出 w_i, 这是比较繁锁的。为了克服这一点, 可采用Logistic 回归分析法, 对死亡率 p_i 进行Logistic 变换:

$$z_i=\ln\frac{p_i}{1-p_i}, \qquad p_i=\frac{1}{1+e^{-z_i}} \tag{10}$$

显然当 $-\infty<z<+\infty$ 时, p 为"S"型曲线, 且 $p=\dfrac{1}{2}$ 时 $z=0$。这样可用回归方程 $\hat{z}=c_0+cx$ 或 $\hat{x}=d_0+dz$ 进行 m 及其方差的估计。由于在 x_i 处 z_i 的方差近似等于$\dfrac{1}{n_ip_i(1-p_i)}$, 因而取权重 $l_i=n_ip_i(1-p_i)$ 对 c_0, c 或 d_0, d 进行加权最小二乘估计, 其结果为:

$$\begin{cases} c=\dfrac{sp}{ss_x}, \qquad c_0=\overline{z}-c\overline{x} \\ \hat{\sigma}^2=\dfrac{1}{\sum l_i-2}\left(ss_z-\dfrac{sp^2}{ss_x}\right) \\ \hat{m}=-\dfrac{c_0}{c} \\ S_z^2=\left[\dfrac{1}{\sum l_i}+\dfrac{(\hat{m}-\overline{x})^2}{ss_x}\right]\hat{\sigma}^2 \\ \mathrm{LD}_{50}\pm S_{\mathrm{LD}_{50}}=10^{\hat{m}}\pm 10^{S_{\hat{z}}} \end{cases} \tag{11}$$

或

$$\begin{cases} d=\dfrac{sp}{ss_z}, \qquad d_0=\overline{x}-d\overline{z} \\ \hat{\sigma}^2=\dfrac{1}{\sum l_i-2}\left(ss_x-\dfrac{sp^2}{ss_z}\right) \\ \hat{m}=d_0 \\ S_{\hat{m}}^2=\left[\dfrac{1}{\sum l_i}+\dfrac{\overline{z}^2}{ss_z}\right]\hat{\sigma}^2 \\ \mathrm{LD}_{50}\pm S_{\mathrm{LD}_{50}}=10^{\hat{m}}\pm 10^{S_{\hat{z}}} \end{cases} \tag{12}$$

其中, $\overline{x}=\dfrac{\sum l_ix_i}{\sum l_i}$, $\overline{z}=\dfrac{\sum l_iz_i}{\sum l_i}$, $ss_x=\sum l_ix_i^2-\dfrac{(\sum l_ix_i)^2}{\sum l_i}$, $ss_z=\sum l_iz_i^2-\dfrac{(\sum l_iz_i)^2}{\sum l_i}$, $sp=\sum l_ix_iz_i-\dfrac{(\sum l_ix_i)(\sum l_iz_i)}{\sum l_i}$。

回归方程的检验仍然用 $r=\dfrac{sp}{\sqrt{ss_xss_z}}$进行。

4 实例与讨论

新、旧机率值法的区别在于机率值 y 和剂量 x 谁作自变量, 新机率值法认为, 估计的是 m, 故 x 应为因变量。Logistic 回归的 2 种方法亦有类似的区别。

机率值法和Logisitic 回归法的主要区别在于二者分别拟合了 2 种"S"型曲线:

$$p_y=\frac{1}{\sqrt{2\pi}}\int_{\infty}e^{-\frac{u^2}{2}}\mathrm{d}u \qquad p_z=\frac{1}{1+e^{-z}}=\frac{1}{1+e^{-c_0u}} \tag{13}$$

二者的共同点是: ① $\lim\limits_{u\to-\infty}p_y=\lim\limits_{u\to-\infty}p_z=0$, $\lim\limits_{u\to+\infty}p_y=\lim\limits_{u\to+\infty}p_z=1$; ②当 $u=0$ 时, $p_y=p_z=\dfrac{1}{2}$。二者的不同在于 p_y 对称而 p_z 不对称。然而 m 的估计是在 $p_y=$

$p_z = \frac{1}{2}$处实现的, 即在 $u = 0$ 邻域实现的。当 u 在 $[-0.1, 0.1]$时, p_y 与 p_z 的差距在 1/1000 左右, 故 Logistic 回归法计算简单而且实用。下面的实例[2]可以说明 2 种方法的一致性。

[例][2]辛硫磷对玉米螟 5 龄幼虫毒力测定分析的部分计算结果见表 1。

表 1　辛硫磷对玉米螟 5 龄幼虫毒力统计分析($n_i = 120$)

Table 1　The toxicity statistic and analysis of phoxim to Ostrinia furnacalis (5 larva) ($n_i = 120$)

λ_i	x_i	$\hat{p}_i$(%)	y_i	$n_i w_i$	$l_i = n_i p_i(1-p_i)$	z_i
6.522	0.814 4	7.6	3.567 5	39.12	8.426 88	-2.498 0
8.696	0.939 3	27.1	4.390 2	62.76	23.707 08	-0.989 6
13.043	1.115 4	51.7	5.042 6	75.84	29.965 32	0.068 0
17.391	1.240 3	81.4	5.892 7	62.76	18.168 48	1.476 2
26.087	1.416 4	93.2	6.490 4	28.92	7.605 12	2.617 8
$\sum$				269.40	87.872 88	

据表 1 得如下计算结果:

$l_{xx} = 8.943\ 3$, $l_{yy} = 217.806\ 8$, $l_{xy} = 43.768\ 0$,

$\bar{x} = 1.092\ 1$, $\bar{y} = 5.029\ 9$, $r = 0.991\ 7^{**}$,

$ss_x = 2.418\ 39$, $ss_z = 167.441\ 84$, $sp = 19.632\ 68$,

$\bar{x} = 1.090\ 9$, $\bar{z} = 0.048\ 43$, $r = 0.975\ 6^{**}$,

机率值法的分析结果为: $a = 4.894\ 1$, $a_0 = -0.314\ 9$, $\hat{m} = 1.086\ 0$, $s_{\hat{y}} = 0.012\ 46$, $LD_{50} = 12.189\ 9$, $S_{LD_{50}} = 1.029\ 1$。

新机率值的分析结果为: $b = 0.201\ 0$, $b_0 = 0.081\ 3$, $\hat{m} = 1.086\ 0$, $s_{\hat{m}} = 0.012\ 3$, $LD_{50} = 12.189\ 9$, $S_{LD_{50}} = 1.028\ 6$.

Logistic 回归的分析结果 1: $c = 8.118\ 08$, $c_0 = -8.807\ 58$, $\hat{m} = 1.084\ 9$, $\hat{\sigma}^2 = 0.093\ 89$, $s_{\hat{z}} = 0.032\ 71$, $LD_{50} = 12.159\ 1$, $S_{LD_{50}} = 1.078\ 2$.

Logistic 回归的分析结果 2: $d = 0.117\ 25$, $d_0 = 1.085\ 22$, $\hat{m} = 1.085\ 22$, $\hat{\sigma}^2 = 0.001\ 356$, $s_{\hat{m}} = 0.003\ 93$, $LD_{50} = 12.168\ 1$, $S_{LD_{50}} = 1.009\ 1$.

由以上计算结果可知, 由于机率值法"S"型曲线的对称性, 使 x 与 y 的相关性较 Logistic 分析法中 x 与 z 的相关性好(0.991 7> 0.975 6)。在 LD_{50} 的估计中, 差异在千分位上, 精确度是一致的, 因而毒力测定的 Logistic 回归分析法在 m 估计的准确度上与机率值法是一致的, 而且计算简单, 是值得使用的一种毒力测定分析方法。

[参考文献]

[1]　张宗炳. 昆虫毒理学[M]. 北京: 科学出版社, 1959. 695-705.

[2]　慕立义. 植物化学保护研究方法[M]. 北京: 中国农业出版社, 1997. 39-45.

[3]　卢恩双, 郭满才, 李　刚, 等. 对毒力测定机率值分析法的新见解[J]. 西北农林科技大学学报(自然科学版), 2001, (3): 104-106.

[4]　袁志发, 周静芋. 多元统计分析[M]. 北京: 科学出版社, 2002. 151-157.

Research on the statistic analysis method for toxicity

YUAN Zhi-fa[1], LU En-shuang[1], GUO Man-Cai[1], Li Gang[2]

(1 College of Life Science; 2 Research and Development Center of Biorational pesticide, Northwest Sci-Tech University of Agriculture and Forestry, YangLing, Shaanxi 712100, China)

Abstract: By study of the probit analysis method for toxicity, two methods based on the Logistic regression have been set up, the comparison between two methods have also been given. The results show that the Logistic regression is not also simple in calculation but also has the same result as the probit value analysis method.

Key words: toxicity; probit value; Logistic regression; half dead dosage; dosage leading to moderate mortality