

同型不同批燃爆产品的感度分布建模^{*}

田玉斌^{1,2}, 李国英², 蔡瑞娇¹

(1. 北京理工大学, 北京 100081;

2. 中国科学院数学与系统科学研究院, 北京 100080)

摘要: 针对具有感度属性的同型不同批燃爆产品的感度分布建模问题, 应用拟合优度的 χ^2 检验法, 给出了产品感度分布的建模方法; 试验数据的同总体检验方法; 感度分布拟合优度检验的条件与方法。应用给出的方法对五批[#]14 针刺火帽的感度分布进行了建模, 实例表明: 同型不同批产品可能具有相同的感度分布, 当对这些批产品的感度分布建模时, 充分利用所有批产品提供的信息是可行的。

关键词: 爆炸力学; 建模; 感度分布; 检验法; 同总体检验; 拟合优度检验

中图分类号: O389; TQ560.72 国标学科代码: 130.3599 文献标志码: A

1 引言

具有感度属性的燃爆产品是武器、弹药及其它燃烧、爆炸系统的关键元件, 这些产品的安全性与可靠性直接关系到武器系统性能的正确发挥和人员安全, 因此对其安全性和可靠性的要求很高。如果简单利用成败型试验数据评估这类产品的可靠性, 所需要的样本量很大。为了有效地减少样本量, 需要利用感度分布的模型信息^[1]。感度分布模型包括分布类型和分布参数两部分。如果错误地选择感度分布类型, 将给评估带来很大的误差^[2], 使这类燃爆产品的使用存在潜在的危害, 并可导致意外事故。

在我国, 有关感度分布建模主要存在着以下几方面的问题:

(1) 在过去很长一个时期, 通常应用几千乃至上万发同型号同一批次产品的试验数据, 拟合产品的感度分布, 如文献[3]使用了同一批次的 11450 发 D6-乙炮弹底火拟合产品的感度分布。生产厂和用户积累的试验数据常常是同型不同批产品的试验数据, 每一批抽检的产品通常不大于 500 发(高价值弹药用产品, 抽检量更少), 所以按照同型同批的抽样方式拟合产品的感度分布较难在实用中推广。

(2) 有的直接将检验合格的同型不同批产品的试验数据进行累加, 并基于该累加数据拟合产品的感度分布。这一方法的缺点是会把服从和不服从同一总体分布的试验数据混合, 将导致感度曲线具有多峰现象, 拟合的感度曲线不能正确地反映产品的感度性能。

(3) 在利用 χ^2 检验公式拟合感度分布时, 通常不谈及试验数据应满足的条件, 从而没有正确地应用 χ^2 检验公式。

针对以上问题, 本文中首先给出了同型不同批产品的试验数据是否服从同一总体的检验方法和公式, 在此基础上给出了感度分布拟合优度检验的条件与方法, 然后给出了感度分布建模的步骤, 并对五批[#]14 针刺火帽的感度分布进行了建模。

2 感度分布的检验

2.1 试验数据的同总体检验

对于合格的同型产品, 可以认为各批产品的感度分布具有相同的类型 $F(y, \theta)$, 不同批产品的感度分布参数可能相同, 也可能不同。如果两批产品的感度分布的参数相同, 那么这两批产品具有相同的感

* 收稿日期: 2002-04-01; 修回日期: 2002-06-11

基金项目: 国家十五计划项目(kkx010304); 国家自然科学基金项目(19631040)

作者简介: 田玉斌(1966—), 女, 博士, 副教授。

度分布,即可认为两批产品的试验数据服从同一分布总体。

对于 m 批同型产品,假设各批产品的感度分布为 $F(y, \theta_i)$, 其中 $F(y, \theta_i)$ 的函数形式未知, $\theta_i = (\theta_{i1}, \theta_{i2}, \dots, \theta_{ir})$ 为 r ($r \leq 3$) 维未知参数, $i = 1, 2, \dots, m$ 。各批产品均在固定的刺激水平 y_1, y_2, \dots, y_k 处独立做刺激试验, 试验数据如下

$$\begin{matrix} & y_1 & y_2 & \cdots & y_k \\ (n_{11}, s_{11}) & (n_{12}, s_{12}) & \cdots & (n_{1k}, s_{1k}) \\ \vdots & \vdots & \vdots & \vdots \\ (n_{m1}, s_{m1}) & (n_{m2}, s_{m2}) & \cdots & (n_{mk}, s_{mk}) \end{matrix} \quad (1)$$

式中: k, m 均大于 3, n_{ij} 为第 i 批产品在刺激水平 y_j 处的试验量, s_{ij} 为 n_{ij} 次试验中的响应数。

为了检验 m 批同型产品有同一个感度分布, 只需检验 $\theta_1 = \theta_2 = \dots = \theta_m$, 从而只需检验

$$p_{11} = p_{21} = \dots = p_{m1}, \quad \dots, \quad p_{1k} = p_{2k} = \dots = p_{mk}$$

式中: $p_{ij} = F(y_j, \theta_i)$ 为第 i 批产品在刺激水平 y_j 下的响应概率。下面给出基于数据 (1) 的检验公式。

定理 1 在数据 (1) 下, 若

$$H_0: p_{11} = p_{21} = \dots = p_{m1}, \quad \dots, \quad p_{1k} = p_{2k} = \dots = p_{mk}$$

成立, 则当 n_{ij} 中的最小量趋于无穷大时, 检验统计量

$$\begin{aligned} Z = & \sum_{j=1}^k \sum_{i=1}^m \frac{\left[s_{ij} - n_{ij} \cdot \frac{s_{1j} + s_{2j} + \dots + s_{mj}}{n_{1j} + \dots + n_{mj}} \right]^2}{n_{ij} \cdot \frac{s_{1j} + \dots + s_{mj}}{n_{1j} + \dots + n_{mj}}} + \\ & \sum_{j=1}^k \sum_{i=1}^m \frac{\left[n_{ij} - s_{ij} - n_{ij} \cdot \frac{(n_{1j} + \dots + n_{mj}) - (s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}} \right]^2}{n_{ij} \cdot \frac{(n_{1j} + \dots + n_{mj}) - (s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}}} \end{aligned} \quad (2)$$

渐近于自由度为 $k(m-1)$ 的 χ^2 分布, 即 Z 渐近于 $\chi^2(k(m-1))$ 。

对于给定的置信水平 $1 - \alpha$, 查表得 $\chi^2(k(m-1))$ 分布的上侧 α 分位数 $\chi^2_{\alpha}(k(m-1))$ 。由定理 1, 如果 Z 大于 $\chi^2(k(m-1))$, 则拒绝假设, 即不能认为这 m 批产品都服从同一总体分布; 否则, 接受假设, 即认为这 m 批产品服从同一总体分布。

2.2 感度分布的拟合优度检验

经过定理 1 检验, 如果试验数据 (1) 来自同一总体, 可将数据进行如下累加

$$\begin{cases} n_j = \sum_{i=1}^m n_{ij} \\ s_j = \sum_{i=1}^m s_{ij} \end{cases} \quad 1 \leq j \leq k$$

得到新的数据

$$\begin{matrix} y_1 & y_2 & \cdots & y_k \\ n_1 & n_2 & \cdots & n_k \\ s_1 & s_2 & \cdots & s_k \end{matrix} \quad (3)$$

式中: n_j 为刺激水平 y_j 下的所有试验发数; s_j 为其中的响应数。基于上述合并后的数据, 可以得到产品在刺激水平 y_j 处响应率的估计

$$p_j^* = \frac{s_j}{n_j} \quad 1 \leq j \leq k$$

合并数据后, 总试验量 $n = \sum_{i=1}^k n_i$ 一般都比较大, 可以利用数据 (3) 进行感度分布类型的检验, 统计学中称为拟合优度检验。具体做法是: 选定一个分布类型, 记为 $G(y, \theta)$ 。这里 $G(y, \theta)$ 的形式已知, r

维参数向量 $\theta \in \Theta$ 未知。现欲考察数据式(3)的分布类型是否为 $G(y, \theta)$ 。以下给出分布类型 $G(y, \theta)$ 的拟合优度检验。

定理 2^[4] 记 $p_i = G(y, \theta)$, 若 ① Θ 是 R^r 的开子集; ② $\partial p_i / \partial \theta_j, 1 \leq i \leq k, j = 1, \dots, r$ 在 Θ 上连续; ③ $\forall \theta \in \Theta$, 矩阵 $\begin{vmatrix} \partial p_1 / \partial \theta_1 & \partial p_1 / \partial \theta_2 & \dots & \partial p_1 / \partial \theta_r \\ \vdots & \vdots & \vdots & \vdots \\ \partial p_k / \partial \theta_1 & \partial p_k / \partial \theta_2 & \dots & \partial p_k / \partial \theta_r \end{vmatrix}$ 的秩为 r ; ④ $\lim \frac{n_i}{n} = \alpha_i > 0, 1 \leq i \leq k$; ⑤ 基于数据式(3)的对数似然函数 $l(\theta)$ 的极小值在有限区域内达到。 $l(\theta)$ 的定义为

$$l(\theta) = \log \left(\prod_{j=1}^k C_{n_j}^{s_j} (G(y_j, \theta))^{s_j} (1 - G(y_j, \theta))^{n_j - s_j} \right) \quad (4)$$

则有

$$\chi^2(\hat{\theta}) = \sum_{j=1}^k \frac{n_j (p_j^* - p_j(\hat{\theta}))^2}{p_j(\hat{\theta})(1 - p_j(\hat{\theta}))} \quad (5)$$

的渐近分布为 $\chi^2(n-r), n \rightarrow \infty$ 。式中: $\hat{\theta}$ 为 θ 的极大似然估计, $p_j(\hat{\theta}) = G(y_j, \hat{\theta})$ 为分布类型 $G(y, \hat{\theta})$ 下, 产品在刺激水平 $y_j (1 < j \leq k)$ 处的响应概率。

通常感度分布 $G(y, \theta)$ 为正态分布、对数正态分布、Logistic 分布、对数 Logistic 分布或威布尔分布。对于这些感度分布, 定理 2 的条件①、②均满足; 当适当选取刺激水平 y_1, y_2, \dots, y_k 时, 条件③也容易满足; 若各个刺激水平下的试验量占总试验量的比例 n_i/n 始终为正常数 $\alpha_i, 1 \leq i \leq k$, 条件④成立; 在数据(3)有混合区间^[1]存在时, 条件⑤成立。此时利用以上拟合优度检验, 当给定置信水平 $1 - \alpha$ 时, 若 $\chi^2(\hat{\theta}) > \chi_\alpha^2(n-r)$, 则认为感度分布类型不是 $G(y, \theta)$; 否则, 可以认为感度分布为 $G(y, \hat{\theta})$ 。

3 感度分布的建模

利用上一节的结果给出感度分布的建模步骤, 然后对五批[#]14 针刺火帽的感度分布进行建模。

3.1 感度分布建模的步骤

①对感度数据(1)进行同总体检验, 若接受检验, 则继续②; 否则, 继续③;

②将同总体数据累加成数据(3)的形式, 转④;

③剔除第 i^* 批数据, 令 $m = m - 1$, 转①;

④利用数据(3), 应用概率单位法^[5]就正态分布、对数正态分布、Logistic 分布、对数 Logistic 分布和威布尔分布分别拟合感度曲线, 并与散点图 $\{(y_i, p_i^*)\}$ 相比较, 初选出最接近散点图的分布做为该产品的感度分布类型;

⑤对于初选出的感度分布类型, 利用似然函数式(4)求参数的最大似然估计, 并利用公式(5)做感度分布的拟合优度检验。

3.2 [#]14 针刺火帽感度分布建模

表 1 给出 5 批合格的[#]14 针刺火帽在刺激水平分别为 0.5、1.0、1.5、2.0、2.5、3.0 cm 的试验数据, 以下利用这 5 批产品的数据进行感度分布的建模(检验水平 $\alpha = 0.05$)。

表 1 5 批合格的[#]14 针刺火帽的试验数据

Table 1 The sensitivity data for #14 cap collected from 5 batches

批次	0.5 cm		1.0 cm		1.5 cm		2.0 cm		2.5 cm		3.0 cm	
	试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率
1	100	0	50	0.20	50	0.98	50	0.98	50	1.00	100	1.00
2	100	0	50	0.24	50	0.80	50	0.98	50	1.00	100	1.00
3	100	0	50	0.22	50	0.86	50	1.00	50	1.00	100	1.00
4	100	0	50	0.22	50	0.90	50	1.00	50	1.00	100	1.00
5	100	0	50	0.28	50	0.92	50	1.00	50	1.00	100	1.00

利用定理 1 进行检验, 此时 $\chi^2_{0.05}(24)=36.415, Z=13.4407$, 故接受假设, 认为这 5 批数据来自同一总体。将这 5 批产品的感度数据累加为式 (3) 的形式, 得到如表 2 的数据。

表 2 5 批 #14 针刺火帽的累加感度数据

Table 2 The sensitivity data for 5 batches caps of #14

0.5 cm		1.0 cm		1.5 cm		2.0 cm		2.5 cm		3.0 cm	
试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率	试验量	发火频率
500	1.00	250	0.232	250	0.892	250	0.992	250	1.00	250	1.00

对表 2 的累加数据, 应用概率单位法拟合感度分布并与散点图进行比较, 初选出最接近散点图的分布。图 1~3 分别给出了与散点图拟合最好, 较好和最差的三种感度分布类型。图中 p 为发火概率, x 为落高。

经过比较可知, 对数 Logistic 分布最接近散点图。因此, 选定对数 Logistic 分布作为感度分布类型, 然后应用似然函数式 (4), 求出参数的最大似然估计 $\hat{\mu}=0.145346, \hat{\sigma}=0.117693$ 。由式 (5) 计算可得 $\chi^2(\hat{\theta})=7.3088$, 查表得 $\chi^2_{0.05}(4)=9.488$, 所以接受对数 Logistic 分布。从图 4 来看, 当参数为 $\hat{\mu}=0.145346, \hat{\sigma}=0.117693$ 时, 对数 Logistic 分布的拟合效果是非常好的。

若假设感度分布为其他四种类型, 由式 (5) 计算可得 $\chi^2(\hat{\theta})$ 值均大于 $\chi^2_{0.05}(4)=9.488$, 因此拒绝其他四种分布作为 #14 针刺火帽的感度分布。

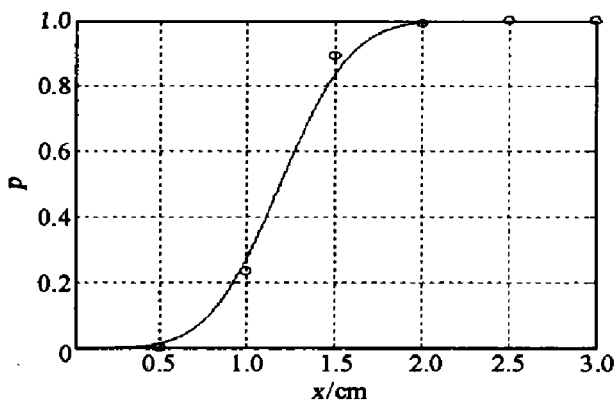


图 1 拟合的正态分布与响应频率的比较
Fig. 1 Comparison between the fitted normal distribution and the responding frequency

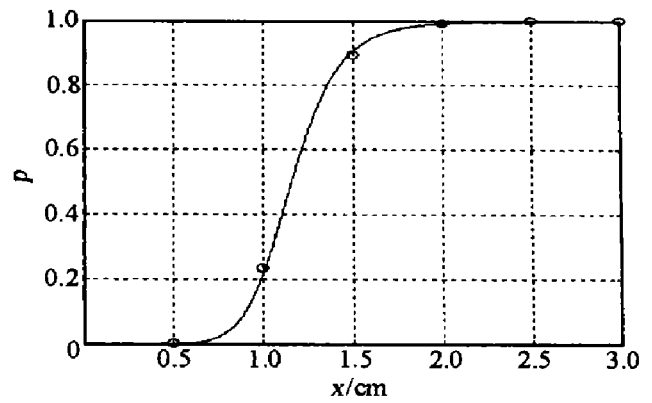


图 2 拟合的对数 Logistic 分布与响应频率的比较
Fig. 2 Comparison between the fitted Log-logistic distribution and the responding frequency

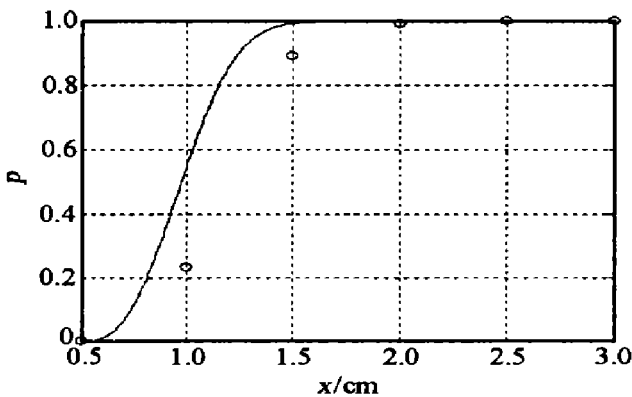


图 3 拟合的 Weibull 分布与响应频率的比较
Fig. 3 Comparison between the fitted Weibull distribution and the responding frequency

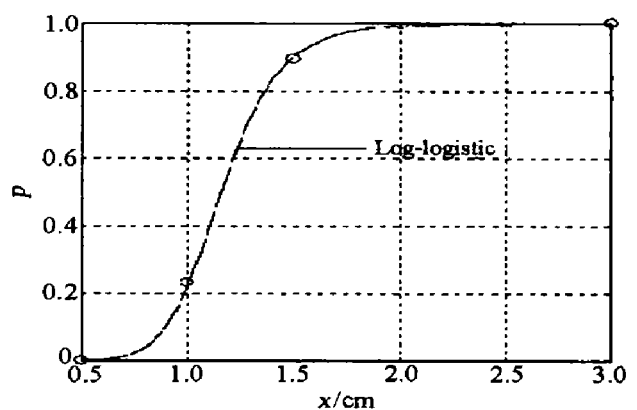


图 4 估计出的对数 Logistic 分布与响应频率的比较
Fig. 4 Comparison between the estimated Log-logistic distribution and the responding frequency

4 结 论

由定理 1 和对 5 批[#]14 针刺火帽感度分布的拟合结果可知, 来自同型号不同批次的产品可能具有完全相同的感度分布。在拟合这些批产品的感度分布时, 充分利用所有批产品提供的信息是可行的。当然, 首先需要通过定理 1 的方法检验哪些批产品具有相同的感度分布。

另外, 可以推断, 本文中讨论的感度分布建模方法也适用于材料强度阈值分布的建模。同总体检验公式也适用于引信等产品的成败型数据是否服从同一总体的检验。

参考文献:

[1] Wu C F J. Efficient sequential designs with binary data[J]. Journal of the American Statistical Association, 1985(8): 974—984.

[2] 钟海芳, 田玉斌, 蔡瑞娇. 感度变量分布类型[J]. 火工品, 1998(3): 1—6.
ZHONG Hai-fang, TIAN Yu-bin, CAI Rui-jiao. Simulation of the sensitivity response curve[J]. Initiators and Pyrotechnics, 1998(3): 1—6.

[3] 严楠. 感度试验设计方法的若干研究[D]. 北京: 北京理工大学, 1996.

[4] Santner Y J, Duffy D E. The statistical analysis of discrete data[M]. New York: Springer-Verlag Inc, 1989.

[5] Rao C R. 线性统计推断及其应用[M]. 张燮. 北京: 科学出版社, 1989.

附录: 定理 1 的证明

由于各个刺激水平处的试验是独立进行的, 先考虑以下检验

$$H_0: p_{11} = p_{21} \cdots = p_{m1} = p_1$$

考虑统计量

$$U = \sum_{i=1}^m \frac{\left[s_{i1} - \frac{n_{i1}(s_{11} + \cdots + s_{m1})}{n_{11} + \cdots + n_{m1}} \right]^2}{\frac{n_{i1}(s_{11} + \cdots + s_{m1})}{n_{11} + \cdots + n_{m1}}} + \sum_{i=1}^m \frac{\left[n_{i1} - s_{i1} - n_{i1} \cdot \frac{(n_{11} + \cdots + n_{m1}) - (s_{11} + \cdots + s_{m1})}{n_{11} + \cdots + n_{m1}} \right]^2}{n_{i1} \cdot \frac{(n_{11} + \cdots + n_{m1}) - (s_{11} + \cdots + s_{m1})}{n_{11} + \cdots + n_{m1}}}$$

$$\hat{p}_{i1} = \frac{s_{i1}}{n_{i1}}, \quad p_1^* = \frac{\sum_{i=1}^m s_{i1}}{\sum_{i=1}^m n_{i1}}, \quad \mathbf{b} \triangleq \begin{pmatrix} \sqrt{\frac{(s_{11} - n_{11}p_1^*)^2}{n_{11}p_1^*(1-p_1^*)}} \\ \vdots \\ \sqrt{\frac{(s_{m1} - n_{m1}p_1^*)^2}{n_{m1}p_1^*(1-p_1^*)}} \end{pmatrix}$$

式中: \hat{p}_{i1} 为第 i 批产品在刺激水平 y_1 处的响应频率; p_1^* 为所有 m 批产品在刺激水平 y_1 处的响应频率; \mathbf{b} 为 m 维向量。

$$\text{令 } \mathbf{V} \triangleq \begin{pmatrix} \frac{n_{11}}{n_{11} + \cdots + n_{m1}} & \frac{\sqrt{n_{11}n_{21}}}{n_{11} + \cdots + n_{m1}} & \cdots & \frac{\sqrt{n_{11}n_{m1}}}{n_{11} + \cdots + n_{m1}} \\ \frac{\sqrt{n_{11}n_{21}}}{n_{11} + \cdots + n_{m1}} & \frac{n_{21}}{n_{11} + \cdots + n_{m1}} & \cdots & \frac{\sqrt{n_{21}n_{m1}}}{n_{11} + \cdots + n_{m1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sqrt{n_{11}n_{m1}}}{n_{11} + \cdots + n_{m1}} & \frac{\sqrt{n_{21}n_{m1}}}{n_{11} + \cdots + n_{m1}} & \cdots & \frac{n_{m1}}{n_{11} + \cdots + n_{m1}} \end{pmatrix}, \quad \mathbf{J} \triangleq \begin{pmatrix} \frac{\sqrt{n_{11}}(\hat{p}_{11} - p_1)}{\sqrt{p_1^*(1-p_1^*)}} \\ \vdots \\ \frac{\sqrt{n_{m1}}(\hat{p}_{m1} - p_1)}{\sqrt{p_1^*(1-p_1^*)}} \end{pmatrix}$$

式中: \mathbf{V} 为 $m \times m$ 矩阵, \mathbf{J} 为 m 维向量。

由于在零假设 H_0 成立条件下, 当 $n^* = \min(n_{11}, n_{21}, \cdots, n_{m1}) \rightarrow \infty$ 时, \mathbf{J} 渐近正态分布 $N_m(0, \mathbf{I})$; 又因为 $(\mathbf{I} - \mathbf{V})$ 为对称幂等阵, 且 $\text{tr}(\mathbf{I} - \mathbf{V}) = m - 1$ 。所以

$$U = \mathbf{b}'\mathbf{b} = \mathbf{J}'(\mathbf{I} - \mathbf{V})'(\mathbf{I} - \mathbf{V})\mathbf{J} = \mathbf{J}'(\mathbf{I} - \mathbf{V})\mathbf{J}$$

渐近于 $\chi^2(m - 1)$ 分布^[5]。因为各刺激量下的试验是独立的, 故当零假设 H_0 成立时

$$\sum_{j=1}^k \sum_{i=1}^m \frac{\left[s_{ij} - \frac{n_{ij}(s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}} \right]^2}{\frac{n_{ij}(s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}}} + \sum_{j=1}^k \sum_{i=1}^m \frac{\left[n_{ij} - s_{ij} - n_{ij} \cdot \frac{(n_{1j} + \dots + n_{mj}) - (s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}} \right]^2}{n_{ij} \cdot \frac{(n_{1j} + \dots + n_{mj}) - (s_{1j} + \dots + s_{mj})}{n_{1j} + \dots + n_{mj}}}$$

渐进于 $\chi^2(k(m-1))$ 分布。

Modeling of response curves for explosive elements of same type but different batches

TIAN YU-bin^{1, 2}, LI Guo-ying², CAI Rui-jiao¹

(1. *Beijing Institute of Technology, Beijing 100081, China;*

2. Institute of Mathematics and System Science, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: For modeling the response curves of explosive elements with sensitivity property and of same type but different batches, a method has been proposed in this paper with the approach of χ^2 goodness-of-fit test. The method consists of the identical population test and the goodness-of-fit test for sensitivity response curves, and has been applied to fit the response curve of 14 caps of five batches. The theoretical and practical results show that the response curves for explosive elements of same type but different batches may be identical. When modeling the response curves of explosive elements from these batches, it is feasible to use the information provided by all these batches.

Key words: mechanics of explosion; modeling; distribution of sensitivity; test method; identical population testing; goodness-of-fit test