# DETERMINING THE NO-OBSERVED-ADVERSE-EFFECT LEVEL IN CATEGORICAL DATA

Nishiyama, Harutoshi
Department of Mathematics, Kyushu University

Yanagawa, Takashi
Department of Mathematics, Kyushu University

https://doi.org/10.5109/13439

出版情報：Bulletin of informatics and cybernetics. 26 (1/2), pp.141-152, 1994-03. Research Association of Statistical Sciences
バージョン：
権利関係：

*Bulletin of Informatics and Cybernetics, Vol. 26, No. 1∼2, 1994*

# DETERMINING THE NO-OBSERVED-ADVERSE-EFFECT LEVEL IN CATEGORICAL DATA

By

**Harutoshi NISHIYAMA*** and **Takashi YANAGAWA†**

**Abstract**

We discuss the determination of the no-observed-adverse-effect level (NOAEL) from a categorical data. Recently a method which incorporated the order restriction into the multiple testing was proposed in Brown and Erdreich [2]. The test is an exact test and computationally involved for large samples. Therefore we propose an alternative test which is competitive to their test and is easily used for large samples.

*Key words and phrases:categorical data, BLV test, asymptotic distribution, PAVA, hypergeometric distribution, random walks, Williams test.*

## 1. Introduction

An experiment with categorical response data is described by the number of experimental objects at risk ($n_i$), the number of interesting response ($r_i$), and the exposure level ($d_i$), for $i = 0, 1, \cdots, k$ as given in Table 1. The subscript zero refers to control group, making $d_0 = 0$ ; otherwise the dose values are arbitrary, subject to order $0 = d_0 < d_1 < \cdots < d_k$ . The true, but unknown response rate at dose $d_i$ is denoted by $p_i$, $i = 0, 1, \cdots, k$ .

**Table 1.** Categorical response data

| dose | $d_0$ | $d_1$ | $\cdots$ | $d_i$ | $\cdots$ | $d_k$ | total |
|---|---|---|---|---|---|---|---|
| response | $r_0$ | $r_1$ | $\cdots$ | $r_i$ | $\cdots$ | $r_k$ | $r_+$ |
| non-response | | | | | | | |
| total | $n_0$ | $n_1$ | $\cdots$ | $n_i$ | $\cdots$ | $n_k$ | $N$ |

In this paper, it is assumed that the samples are random and mutually independent, and that the number of response $r_i$ at $d_i$ is distributed as binomial distribution B $(n_i, p_i)$ with parameters $n_i$ and $p_i$ for $i = 0, 1, \cdots, k$. It is also assumed to be known *a priori* that the true response rate is nondecreasing as dose increases, *i.e.* $0 \le p_0 \le p_1 \le \cdots \le$

* Department of Mathematics, Kyushu University 33, Fukuoka 812, Japan
**E-mail:** nisiyama@math.sci.kyushu-u.ac.jp

† Department of Mathematics, Kyushu University 33, Fukuoka 812, Japan

$p_k \leq 1$. The purpose of this paper is to consider methods to decide $d_i$ such that $p_0 = p_1 = \cdots = p_i < p_{i+1}$. In application this $d_i$ is often called the no-observed-adverse-effect level(NOAEL).

The methods of multiple comparison such as Dunnett's [3] and Scheffé's [9] may be applied to this problem. Recently, an interesting method was proposed in Brown and Erdreich [2]. The method incorporates the order restriction $0 \leq p_0 \leq p_1 \leq \cdots \leq p_k \leq 1$ into the multiple conditional testing conditioned on all margins. The test, which is an exact test and computationally involved for large samples, is called the Brown-La Vange test(BLV test)(Brown and Erdreich [2]).

We propose in this paper an alternative test which also is exact and furthermore has an asymptotic approximation, and study its characteristic.

The BLV test and a new test are described in section 2. The property of the new test is examined in section 3. In section 4, an asymptotic distribution is obtained for the critical points of the test when the number of sample is large. Also the difference between the asymptotic and the exact distribution is evaluated when the sizes of samples are small. In section 5 we compare the new test with the BLV test, and it is indicated that the proposed test is competitive to the BLV test.

## 2. Testing Procedures

### 2.1. Brown-La Vange test

For the null hypothesis $H_0^{(i)} : p_0 = p_1 = \cdots = p_i$ , the test uses $T_i := \hat{p}_i - \hat{p}_0$ as the test statistics, where $\hat{p}_i$ is the maximum likelihood estimate (m.l.e.) of $p_i$ under the constraint $p_0 \leq p_1 \leq \cdots \leq p_k$ . The m.l.e. of $p_i$ 's under the order restriction are constructed by the Pooled-Adjacent-Violators Algorithm (PAVA)(Robertson, Wright and Dykstra [8]). It is well known that $\hat{p}_i$ may be expressed by the max-min formulas as follows:

$$\hat{p}_i := \max_{0 \leq u \leq i} \min_{i \leq v \leq k} \left( \sum_{j=u}^{v} r_j / \sum_{j=u}^{v} n_j \right) \quad (i = 0, 1, 2, \cdots, k).$$

In the BLV procedure, the null hypothesis $H_0^{(k)} : p_0 = p_1 = \cdots = p_k$ is tested initially. For a specified test size $\alpha_1$ , reject $H_0^{(k)}$ if $T_k$ takes a value greater or equal to $C_k(\alpha_1)$ , where $C_k(\alpha_1)$ is the smallest constant $C$ such that $\Pr[T_k \geq C | r_+] \leq \alpha_1$ when $H_0^{(k)}$ is true. If $H_0^{(k)}$ is not rejected, then the NOAEL takes the value $d_k$ and the test is ended. If $H_0^{(k)}$ is rejected, then $H_0^{(k-1)}$ is tested. For a specified test size $\alpha_2$ , reject $H_0^{(k-1)}$ if $T_{k-1}$ takes a value greater or equal to $C_{k-1}(\alpha_2)$ , where $C_{k-1}(\alpha_2)$ is the smallest constant $C$ such that $\Pr[T_{k-1} \geq C | r_+, T_k \geq C_k(\alpha_1)] \leq \alpha_2$ when $H_0^{(k-1)}$ is true. If $H_0^{(k-1)}$ is not rejected, then the NOAEL takes the value $d_{k-1}$ and the test is ended. If $H_0^{(k-1)}$ is rejected, then $H_0^{(k-2)}$ is similarly tested for a specified test size $\alpha_3$ and so on. If all null hypotheses are rejected, then the NOAEL takes the value $d_0$ and the test is ended.

### 2.2. An alternative test

The computation for $C_i(\alpha)$ of the BLV test is involved when the sizes of samples are large. We modify the BLV test so that we may obtain not only an exact, but also an approximate critical point of the test.
Instead of $T_i$ we use

$$M_i := \bar{p}_i - p^*,$$

for testing the $H_0^{(i)}$, where

$$\bar{p}_i := \max_{1 \le u \le i} \left(\sum_{j=u}^{i} r_j / \sum_{j=u}^{i} n_j\right) \quad (i = 1, 2, \cdots, k), \tag{2.1}$$

and

$$p^* := \frac{r_0}{n_0}.$$

The testing procedure of the new test is the same as the BLV test except for the determination of the critical values. The critical value $C_i^*(\alpha)$ of the new test is determined as the smallest constant $C$ such that $\Pr[M_i \ge C | r_+, H_0^{(k)}$ is true$] \le \alpha$ ($i = k, k-1, \cdots, 1$). Note that we use the same $\alpha$ for each stage. Note also that the following inequality holds

$$\alpha \ge \Pr[M_i \ge C | r_+, H_0^{(k)} \text{is true}] \ge \Pr[M_i \ge C | r_+, H_0^{(i)} \text{is true}]$$

for all $C > 0$ and all $i = k, k-1, \cdots, 1$.

We call this test the modified Brown-La Vange test(MBLV test).

## 3. The property of the MBLV test

### 3.1. Type I FWE

Type I FWE(familywise error) is the probability of rejecting at least one true hypotheses. In this problem, supposing $H_0^{(j_0)}$ is the true null hypothesis, Type I FWE is represented in the present set up by

$$\text{Type I FWE} := \Pr[\bigcup_{i=1}^{j_0} \{ \text{ reject } H_0^{(i)} \} | H_0^{(j_0)} \text{ is true}].$$

We may prove the following inequality

$$\text{Type I FWE} \le \Pr[\text{ reject } H_0^{(j_0)} \; | H_0^{(j_0)} \text{ is true}] \le \alpha.$$

Note that this inequality is a special case of the theorem by Marcus, Peritz and Gabriel [5].

### 3.2. Characteristics of the statistics

The statistics $M_i$ in the MBLV testing procedure uses $\bar{p}_i$ as an estimator of $p_i$. In this section we show that we may replace $\bar{p}_i$ with $\hat{p}_i$ in the procedure. Here $\bar{p}_i$ is given in (2.1) and $\hat{p}_i$ is the m.l.e. of $p_i$ under the order restriction $p_0 \le p_1 \le \cdots \le p_k$. For this aim we introduce $\tilde{p}_i$ in THEOREM 3.1 and consider the relationship of $\hat{p}_i$ and $\tilde{p}_i$, and in THEOREM 3.2 we consider that of $\tilde{p}_i$ and $\bar{p}_i$.

THEOREM 3.1. *Let $\tilde{p}_i$ denote the m.l.e. of $p_i$ under the order restriction $p_1 \le p_2 \le \cdots \le p_k$ $(i = 1, 2, \cdots, k)$ where $p_0$ is not included. Then if $\hat{p}_i - (r_0/n_0) > 0$ or $\tilde{p}_i - (r_0/n_0) > 0$, $\hat{p}_i$ is equal to $\tilde{p}_i$ $(i = 1, 2, \cdots, k)$.*

PROOF. It is sufficient to show the theorem when $\hat{p}_i - (r_0/n_0) > 0$ since $\hat{p}_i \ge \tilde{p}_i$ for $\forall i \in \{1, 2, \cdots, k\}$.

Let $(A_1, A_2, \cdots, A_r)$ be the solution block (Barlow, Bartholomew, Bremner and Brunk [1]) such that on each $A_i$ the restricted m.l.e.'s are constant. It is clear that $0 \in A_1$. Furthermore, if $i$ belongs to $A_j$, then

$$\hat{p}_i = \mathrm{Av}(A_j),$$

where

$$\mathrm{Av}(A_j) := \frac{\sum_{l \in A_j} r_l}{\sum_{l \in A_j} n_l}.$$

First we prove the theorem when $i = 1$. If $1 \in A_1$, then

$$\hat{p}_1 = \hat{p}_0 = Av(A_1) = \min_{0 \le v \le k} \frac{\sum_{j=0}^{v} r_j}{\sum_{j=0}^{v} n_j} \le \frac{r_0}{n_0}.$$

This conflicts with $\hat{p}_1 - (r_0/n_0) > 0$. Thus $1 \notin A_1$, namely $\hat{p}_0 = (r_0/n_0)$ and $\hat{p}_1 = \tilde{p}_1$. Next we prove the theorem for $i > 1$. In general, if $\hat{p}_{i_1} > \hat{p}_{i_2}$, there exist two integers $j_1$ and $j_2$ such that $1 \le j_2 < j_1 \le r$, $\hat{p}_{i_1} = \mathrm{Av}(A_{j_1})$ and $\hat{p}_{i_2} = \mathrm{Av}(A_{j_2})$. Therefore $\hat{p}_{i_1}$ does not contain $r_i/n_i$ for any $i \notin A_{j_1}$. Thus if $\hat{p}_i > (r_0/n_0)$, we have i feasible cases of the location of $r_0/n_0$, namely

$$\hat{p}_i \ge \hat{p}_{i-1} \ge \cdots \ge \hat{p}_2 \ge \hat{p}_1 > \frac{r_0}{n_0} = \hat{p}_0,$$

$$\hat{p}_i \ge \hat{p}_{i-1} \ge \cdots \ge \hat{p}_2 > \frac{r_0}{n_0} \ge \hat{p}_1 \ge \hat{p}_0,$$

$$\vdots$$

or

$$\hat{p}_i > \frac{r_0}{n_0} \ge \hat{p}_{i-1} \ge \cdots \ge \hat{p}_2 \ge \hat{p}_1 \ge \hat{p}_0,$$

but in any case we have $\hat{p}_i = \tilde{p}_i$. □

THEOREM 3.2. *Let $\tilde{p}_i$ and $\bar{p}_i$ be the estimators of $p_i$ defined in* THEOREM 3.1 *and formula* (2.1), *respectively, then we have*

$$\tilde{p}_j > t_j \quad (\text{ for } \forall j = i, \cdots, k) \Longleftrightarrow \bar{p}_j > t_j \quad (\text{ for } \forall j = i, \cdots, k)$$

for $t_i \le t_{i+1} \le \cdots \le t_k$ $(i \in \{1, 2, \cdots, k\})$.

PROOF. First we show ($\Longleftarrow$).
Let $(B_1, B_2, \cdots, B_r)$ be the solution block such that on each $B_i$ the restricted m.l.e.'s are constant. Furthermore, if $i \in B_m$, then

$$\tilde{p}_i = \mathrm{Av}(B_m),$$

where

$$\mathrm{Av}(B_m) := \frac{\sum_{l \in B_m} r_l}{\sum_{l \in B_m} n_l}.$$

Put $B_m = \{h, h+1, \cdots, h+h_1\}$, then

$$\tilde{p}_h = \tilde{p}_{h+1} = \cdots = \tilde{p}_{h+h_1} = \mathrm{Av}(B_m) = \frac{\sum_{l=h}^{h+h_1} r_l}{\sum_{l=h}^{h+h_1} n_l}. \tag{3.1}$$

Furthermore, from

$$\tilde{p}_{h+h_1} = \max_{1 \le u \le h+h_1} \min_{h+h_1 \le v \le k} \frac{\sum_{l=u}^{v} r_l}{\sum_{l=u}^{v} n_l},$$

from the uniqueness of $(B_1, B_2, \cdots, B_r)$ and also from (3.1) we have

$$\tilde{p}_{h+h_1} = \max_{1 \le u \le h+h_1} \frac{\sum_{l=u}^{h+h_1} r_l}{\sum_{l=u}^{h+h_1} n_l} = \bar{p}_{h+h_1}.$$

Therefore

$$\tilde{p}_{h+h_1} = \bar{p}_{h+h_1}. \tag{3.2}$$

Now for $\forall j \in \{i, \cdots, k\}$, there exists $m_1 \in \{1, \cdots, r\}$ such that $j \in B_{m_1}$. Thus

$$\tilde{p}_j = \mathrm{Av}(B_{m_1}).$$

Put

$$j_1 := \max_{l \in B_m} l,$$

then from $j_1 \ge j$ and (3.1),

$$\tilde{p}_j = \tilde{p}_{j_1}.$$

Therefore, since (3.2) and $t_i \le \cdots \le t_j \le \cdots \le t_{j_1} \le \cdots \le t_k$, it follows that

$$\tilde{p}_j = \tilde{p}_{j_1} = \bar{p}_{j_1} > t_{j_1} \ge t_j.$$

$$\tilde{p}_j > t_j.$$

Therefore ($\Longleftarrow$) is shown.

Next we show ($\Longrightarrow$). Now it follows that

$$\tilde{p}_j = \max_{1 \le u \le j} \min_{j \le v \le k} \frac{\sum_{l=u}^{v} r_l}{\sum_{l=u}^{v} n_l} \le \max_{1 \le u \le j} \frac{\sum_{l=u}^{j} r_l}{\sum_{l=u}^{j} n_l} = \bar{p}_j.$$

Therefore

$$\tilde{p}_j \le \bar{p}_j.$$

From this inequality if $\tilde{p}_j > t_j$ (for $\forall j = i, \cdots, k$), then

$$\bar{p}_j \ge \tilde{p}_j > t_j \text{ (for } \forall j = i, \cdots, k).$$

Therefore ($\Longrightarrow$) is also shown. □

## 4. The critical points

In this section, we consider the computation of getting the critical points $C_i^*(\alpha)$. Section 4.1 is devoted to the exact case. In section 4.2 we consider an asymptotic distribution of $M_i$. In section 4.3 we obtain the critical points when $n_i$ and $r_+$ are large.

### 4.1. Exact case

We first list up all configurations of $(r_0, r_1, \cdots, r_k)$ when $\{n_0, n_1, \cdots, n_k, r_+\}$ is given, and obtain the distribution of $M_i$, by computing the probability of the occurence of each configuration and also computing the corresponding value of $M_i$. Then determine $C_i^*(\alpha)$ as the smallest constant $C$ such that $\Pr[M_i \geq C|r_+, H_0^{(k)}\text{is true}] \leq \alpha$ with the distribution of $M_i$.

As an example, we shall get the critical points $C_i^*(\alpha)$ for $k = 2$, $n_0 = n_1 = n_2 = 20$, $r_+ = 5$, $\alpha = 0.05$. Table 2 lists all configurations of $(r_0, r_1, r_2)$, their occurence probabilities and the values of $M_i$ and $T_i$. The distributions of $M_2$, $M_1$ and also $T_2$, $T_1$ are summarized in Table 3. From Table 3, we determine $C_2^*(0.05) = C_1^*(0.05) = 0.20$.

**Table 2.** The configurations of $(r_0, r_1, r_2)$ for $n_0 = n_1 = n_2 = 20$ and $r_+ = 5$; the probability of each configuration under $p_0 = p_1 = p_2$; and the values of $M_2$, $M_1$, $T_2$ and $T_1$ for each configuration.

| No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $r_2$ | 5 | 4 | 3 | 2 | 1 | 0 | 4 |
| $r_1$ | 0 | 1 | 2 | 3 | 4 | 5 | 0 |
| $r_0$ | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| prob. | 0.0028 | 0.0177 | 0.0397 | 0.0397 | 0.0177 | 0.0028 | 0.0177 |
| $M_2$ | 0.25 | 0.20 | 0.15 | 0.125 | 0.125 | 0.125 | 0.15 |
| $M_1$ | 0.00 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | -0.05 |
| $T_2$ | 0.25 | 0.20 | 0.15 | 0.125 | 0.125 | 0.125 | 0.175 |
| $T_1$ | 0.00 | 0.05 | 0.10 | 0.125 | 0.125 | 0.125 | 0.00 |
| No. | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| $r_2$ | 3 | 2 | 1 | 0 | 3 | 2 | 1 |
| $r_1$ | 1 | 2 | 3 | 4 | 0 | 1 | 2 |
| $r_0$ | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| prob. | 0.0835 | 0.1322 | 0.0835 | 0.0177 | 0.0397 | 0.1322 | 0.1322 |
| $M_2$ | 0.10 | 0.05 | 0.05 | 0.05 | 0.05 | 0.00 | -0.025 |
| $M_1$ | 0.00 | 0.05 | 0.10 | 0.15 | -0.10 | -0.05 | 0.00 |
| $T_2$ | 0.10 | 0.05 | 0.05 | 0.05 | 0.10 | 0.025 | 0.00 |
| $T_1$ | 0.00 | 0.05 | 0.05 | 0.05 | 0.00 | 0.00 | 0.00 |
| No. | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| $r_2$ | 0 | 2 | 1 | 0 | 1 | 0 | 0 |
| $r_1$ | 3 | 0 | 1 | 2 | 0 | 1 | 0 |
| $r_0$ | 2 | 3 | 3 | 3 | 4 | 4 | 5 |
| prob. | 0.0397 | 0.0397 | 0.0835 | 0.0397 | 0.0177 | 0.0177 | 0.0029 |
| $M_2$ | -0.025 | -0.05 | -0.10 | -0.10 | -0.15 | -0.175 | -0.25 |
| $M_1$ | 0.05 | -0.15 | -0.10 | -0.05 | -0.20 | -0.15 | -0.25 |
| $T_2$ | 0.00 | 0.025 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $T_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

**Table 3.** The distribution of $M_2$, $M_1$, $T_2$ and $T_1$ under $p_0 = p_1 = p_2$ for $n_0 = n_1 = n_2 = 20$ and $r_+ = 5$.

| $x$ | 0.00 | 0.05 | 0.10 | 0.125 | 0.15 | 0.20 | 0.25 |
|---|---|---|---|---|---|---|---|
| $\Pr(M_2 \geq x)$ | 0.6269 | 0.4947 | 0.2216 | 0.1381 | 0.0779 | 0.0205 | 0.0028 |
| $x$ | -0.05 | 0.00 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 |
| $\Pr(M_1 \geq x)$ | 0.7988 | 0.6093 | 0.3907 | 0.2011 | 0.0779 | 0.0205 | 0.0028 |
| $x$ | 0.05 | 0.10 | 0.125 | 0.15 | 0.175 | 0.20 | 0.25 |
| $\Pr(T_2 \geq x)$ | 0.4947 | 0.2613 | 0.1381 | 0.0779 | 0.0382 | 0.0205 | 0.0028 |
| $x$ | 0.00 | 0.05 | 0.10 | 0.125 | | | |
| $\Pr(T_1 \geq x)$ | 1.0000 | 0.3510 | 0.0999 | 0.0602 | | | |

## 4.2. Asymptotic distribution

We consider the limiting conditional distribution of $M_i$ under $H_0^{(k)}$, when $n_0 = n_1 = \cdots = n_k = n$, conditioned on $n$ and $r_+$. We first show the two theorems.

THEOREM 4.1. *Limiting conditional distribution of the random vector*

$$\left(\frac{r_0}{n} - \frac{r_1}{n}, \cdots, \frac{r_0}{n} - \frac{r_k}{n}, \frac{r_1}{n} - \frac{r_2}{n}, \cdots, \frac{r_{k-1}}{n} - \frac{r_k}{n}\right) \Big/ \sqrt{\frac{r_+(N - r_+)}{nN(N-1)}}$$

*conditioned on $n$ and $r_+$ is identical to the distribution of*

$$(Z_0 - Z_1, \cdots, Z_0 - Z_k, Z_1 - Z_2, \cdots, Z_{k-1} - Z_k),$$

*where $Z_0$, $Z_1$, $\cdots$, $Z_k$ are random variables which are independently and identically distributed as a standard normal distribution and $N := (k+1)n$.*

PROOF. When $n$ and $r_+$ are given and $H_0^{(k)}$ is true, the conditional distribution of $(r_0, r_1, \cdots, r_k)$ is a multiple hypergeometric distribution. Therefore for any $j$,

$$E(r_j) = r_+ n/N,$$

$$Var(r_j) = r_+ n(N - r_+)(N - n)/\{N^2(N-1)\},$$

and for any $j_1 \neq j$,

$$Cov(r_j, r_{j_1}) = -n^2 r_+(N - r_+)/\{N^2(N-1)\}.$$

Thus we have, for any $j_1 \neq j_2$, $j_1 \neq j_3$ and $j_2 \neq j_3$

$$Var\left(\frac{r_{j_1}}{n} - \frac{r_{j_2}}{n}\right) = 2r_+(N - r_+)/\{nN(N-1)\},$$

$$corr\left(\frac{r_{j_1}}{n} - \frac{r_{j_2}}{n}, \frac{r_{j_1}}{n} - \frac{r_{j_3}}{n}\right) = \frac{1}{2},$$

$$corr\left(\frac{r_{j_1}}{n} - \frac{r_{j_2}}{n}, \frac{r_{j_2}}{n} - \frac{r_{j_3}}{n}\right) = -\frac{1}{2}.$$

Furthermore from the asymptotic normality of the multiple hypergeometric distribution(Plackett [7]),

$$(\frac{r_0}{n}-\frac{r_1}{n},\cdots,\frac{r_0}{n}-\frac{r_k}{n},\frac{r_1}{n}-\frac{r_2}{n},\cdots,\frac{r_{k-1}}{n}-\frac{r_k}{n})$$

is asymptotically distributed as a multiple normal distribution. Thus the proof of the theorem is immediate, since

$$corr(Z_{j_1}-Z_{j_2},Z_{j_1}-Z_{j_3})=\frac{1}{2},$$

$$corr(Z_{j_1}-Z_{j_2},Z_{j_2}-Z_{j_3})=-\frac{1}{2},$$

for any $j_1\neq j_2$, $j_1\neq j_3$ and $j_2\neq j_3$. □

THEOREM 4.2. *Suppose that $X_1,X_2,\cdots,X_n,\cdots$ are independently and identically distributed random variables as a standard normal distribution. Put*

$$U_n:=\max_{1\leq r\leq n}\frac{1}{r}\sum_{i=1}^{r}X_i$$

*and*

$$F(x):=\begin{cases} exp[-\sum_{r=1}^{\infty}\frac{1}{r}\{1-\Phi(xr^{1/2})\}] & x>0\\ 0 & x\leq 0,\end{cases}$$

*where $\Phi$ denotes the normal distribution function. Then*

$$\lim_{n\to\infty}\Pr[U_n\leq x]=F(x).$$

PROOF. Williams [11] showed this theorem for $x>0$ by using the theory of random walks(Feller [4]). We show the theorem for $x\leq 0$ . Since

$$0\leq\lim_{n\to\infty}\Pr[U_n\leq x]\leq\lim_{n\to\infty}\Pr[U_n\leq 0]\ \text{ for }\ x\leq 0,$$

it is sufficient to show that

$$\lim_{n\to\infty}\Pr[U_n\leq 0]=0.$$

Putting $S_r:=\sum_{i=1}^{r}X_i$, we have

$$\lim_{n\to\infty}\Pr[U_n\leq 0]=\lim_{n\to\infty}\Pr[S_1\leq 0,S_2\leq 0,\cdots,S_n\leq 0].$$

Furthermore, putting $W_n:=\max\{S_1,\cdots,S_n\}$, we have

$$\lim_{n\to\infty}\Pr[U_n\leq 0]=\lim_{n\to\infty}\Pr[M_n\leq 0]. \tag{4.1}$$

Now, applying the law of iterated logarithm, namely

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\log\log n}}=1\ a.s.,$$

we have

$$\limsup_{n\to\infty} S_n = +\infty \ a.s.$$

Furthermore, we have

$$\inf_{k\geq n} W_k \geq S_n,$$

thus

$$\liminf_{n\to\infty} W_n \geq \limsup_{n\to\infty} S_n \ a.s.,$$

and

$$W_n \stackrel{a.s.}{\to} +\infty \qquad \text{as } n \to +\infty.$$

Therefore from (4.1)

$$\lim_{n\to\infty} \Pr[U_n \leq 0] = 0 = F(0). \qquad \square$$

From THEOREM 4.1 and 4.2, the following theorem follows.

THEOREM 4.3. *Put*

$$q := \sqrt{\frac{r_+(N-r_+)}{nN(N-1)}}.$$

*Then for $F(t)$ defined in* THEOREM 4.2,

$$\lim_{i,n\to\infty} \Pr[M_i < x | r_+, p_0 = \cdots = p_i = \cdots = p_k] = \int_{-\infty}^{\infty} F(t + \frac{x}{q})\phi(t)dt$$

*,where $\phi$ denotes the normal density function.*

PROOF. From THEOREM 4.1, the limiting conditional distribution of

$$(\frac{r_1}{n} - \frac{r_0}{n}, \frac{r_2}{n} - \frac{r_0}{n}, \cdots, \frac{r_i}{n} - \frac{r_0}{n})/\sqrt{\frac{r_+(N-r_+)}{nN(N-1)}}$$

conditioned on $n$ and $r_+$ is identical to the distribution of

$$(Z_1 - Z_0, Z_2 - Z_0, \cdots, Z_i - Z_0).$$

Therefore the limiting conditional distribution of $M_i/q$ conditioned on $n$ and $r_+$ is identical to the distribution of

$$T := \max_{1\leq u\leq i} \frac{1}{i-u+1} \sum_{j=u}^{i} (Z_j - Z_0)$$

From this and THEOREM 4.2,

$$\begin{aligned}
\lim_{i,n\to\infty} \Pr[M_i < x | r_+, p_0 &= \cdots = p_i = \cdots = p_k] \\
&= \lim_{i\to\infty} \Pr[qT < x] \\
&= \lim_{i\to\infty} \int_{-\infty}^{\infty} \Pr[\max_{1\leq u\leq i} \frac{1}{u} \sum_{j=1}^{u} Z_j < t + \frac{x}{q}]\phi(t)dt \\
&= \int_{-\infty}^{\infty} F(t + \frac{x}{q})\phi(t)dt
\end{aligned} \qquad \square$$

### 4.3. Approximate value of the critical points

Employing the asymptotic distribution of $M_i$ in THEOREM 4.3, we may approximate the exact critical value by the smallest constant $C$ such that

$$\int_{-\infty}^{\infty} F(t+\frac{C}{q})\phi(t)dt \geq 1-\alpha,$$

where $q$ is defined in THEOREM 4.3.

To evaluate this approximation we computed the exact and approximate critical values for test sizes $\alpha = 0.05$ and $\alpha = 0.10$ when $n_0 = n_1 = n_2 = 10$, 20 and 30, and $r_+ = 4$ and 5. The results are summarized in Table 4.

**Table 4.** Critical points for the MBLV test

| | | critical point ($\alpha = 0.05$) | | | | critical point ($\alpha = 0.10$) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $r_+ = 4$ | | $r_+ = 5$ | | $r_+ = 4$ | | $r_+ = 5$ | |
| | | Exa. | Apr. | Exa. | Apr. | Exa. | Apr. | Exa. | Apr. |
| $n_0 = n_1 = n_2 = 10$ | $C_2^*$ | 0.40 | 0.275 | 0.40 | 0.302 | 0.30 | 0.226 | 0.30 | 0.248 |
| | $C_1^*$ | 0.40 | 0.275 | 0.40 | 0.302 | 0.30 | 0.226 | 0.30 | 0.248 |
| $n_0 = n_1 = n_2 = 20$ | $C_2^*$ | 0.20 | 0.142 | 0.20 | 0.157 | 0.15 | 0.117 | 0.15 | 0.129 |
| | $C_1^*$ | 0.20 | 0.142 | 0.20 | 0.157 | 0.15 | 0.117 | 0.15 | 0.129 |
| $n_0 = n_1 = n_2 = 30$ | $C_2^*$ | 0.133 | 0.096 | 0.133 | 0.106 | 0.10 | 0.079 | 0.10 | 0.087 |
| | $C_1^*$ | 0.133 | 0.096 | 0.133 | 0.106 | 0.10 | 0.079 | 0.10 | 0.087 |

Poor approximations might be seen in the tables. However, one must take into account the discreteness of the distribution of $M_i$ in such evaluation. Figure 1 and 2 show the exact and approximate distributions of $M_2$ when $n_0 = n_1 = n_2 = 10$, $r_+ = 4$ and for $n_0 = n_1 = n_2 = 30$, $r_+ = 5$, respectively. The inspection of the figures indicates that the approximations in Table 4 are not so bad when the discreteness is taken into account.

## 5. Comparison of the two tests

We compare the MBLV test with the BLV test when $k = 2$, $n_0 = n_1 = n_2 = 20$, and $r_+ = 5$. Using Table 2 we obtain the NOAEL by means of the MBLV test and BLV test. The results are summarized in Table 5. The table shows that for the MBLV test the configurations No.1 and No.2 select $d_1$ and any others select $d_2$ as the NOAEL at the test size $\alpha = 0.05$; that for the BLV test the configurations No.1, No.2 and No.7 select $d_1$ and any others select $d_2$ at the test size $\alpha_1 = \alpha_2 = 0.05$.

Thus when $\alpha = \alpha_1 = \alpha_2 = 0.05$ it follows from Table 5 that the probability of correct decision by the MBLV test is smaller than the BLV test when $p_0 = p_1 < p_2$; larger than the BLV test when $p_0 = p_1 = p_2$.

This finding comes from the fact that the carrier of the distribution of $T_2$ includes that of the distribution of $M_2$ (*see* Table 3), and that $T_2$ selects the critical points which are closer to the nominal size than $M_2$. This discrepancy would decrease if the sample sizes increase, or if large test sizes are taken. For example, when $\alpha = \alpha_1 = \alpha_2 = 0.10$, Table 5 shows that the configurations No.1, No.2, No.3 and No.7 select $d_1$ and any others select $d_2$ as the NOAEL for both tests.

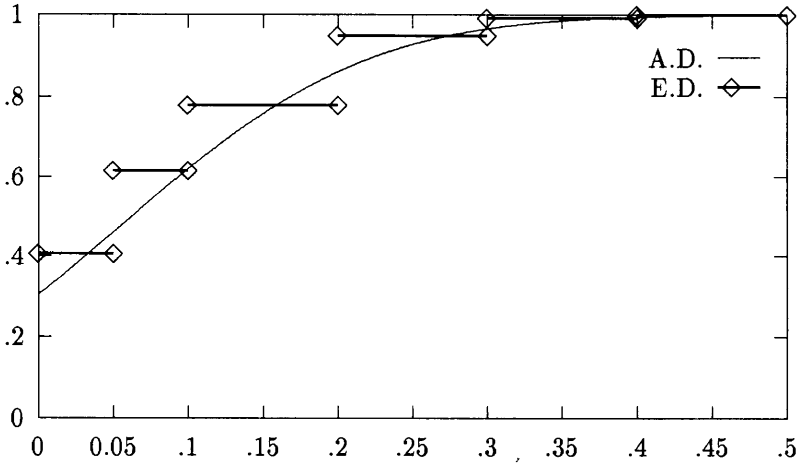


**Figure 1.** The approximate distribution (A.D.) and the exact distribution (E.D.) of $M_2$ for $n_0 = n_1 = n_2 = 10$ and $r_+ = 4$.

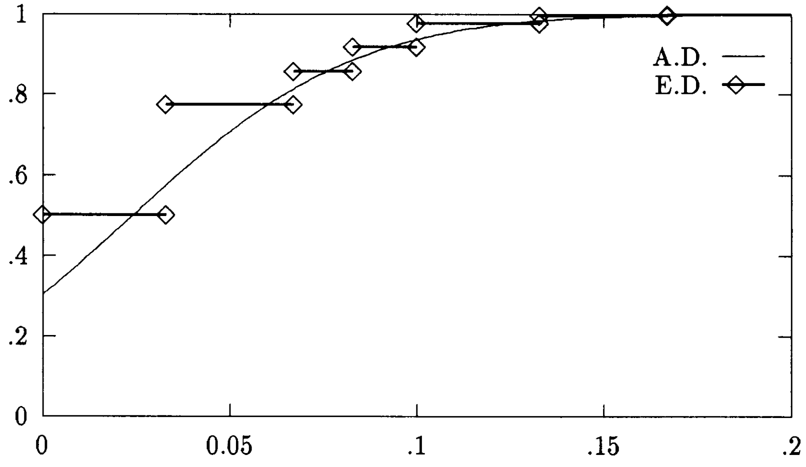


**Figure 2.** The approximate distribution (A.D.) and the exact distribution (E.D.) of $M_2$ for $n_0 = n_1 = n_2 = 30$ and $r_+ = 5$.

**Table 5.** The configuration which select $d_0$, $d_1$ and $d_2$ as the NOAEL when the MBLV test and the BLV test are applied for $n_0 = n_1 = n_2 = 20$, $r_+ = 5$.

| | Test size (critical point) | | NOAEL | | |
|---|---|---|---|---|---|
| MBLV test | $\alpha$ ($C_2^*(\alpha)$) ($C_1^*(\alpha)$) | | $d_0$ | $d_1$ | $d_2$ |
| | 0.05 (0.20) (0.20) | | none | 1,2 | a.o. |
| | 0.10 (0.15) (0.15) | | none | 1,2,3,7 | a.o. |
| BLV test | $\alpha_1$ ($C_2(\alpha_1)$) | $\alpha_2$ ($C_1(\alpha_2)$) | $d_0$ | $d_1$ | $d_2$ |
| | 0.05 (0.175) | 0.05 (0.10) | none | 1,2,7 | a.o. |
| | 0.10 (0.15) | 0.10 (0.125) | none | 1,2,3,7 | a.o. |

a.o.:all the others

When this paper was presented at a symposium, Professor Hirotsu pointed up that the MBLV test and BLV test were categorical data versions of the Williams test [10] and modified Williams test [10]. Marcus [6] conducted a Monte Carlo study, compared the powers of the Williams test and modified Williams test, and found that these tests were competitive (*see* Marcus [6] Table 3(a)). The same would be expected for the powers of the BLV test and MBLV test.

## References

[ 1 ] Barlow,R.E., Bartholomew,D.J., Bremner,J.M. and Brunk,H.D.: *Statistical Inference under Order Restriction.* JOHN WILEY & SONS, (1972).

[ 2 ] Brown,K.G. and Erdreich,L.S.: *Statistical uncertainly in the No-observed- adverse-effect level.* Fundamental and Applied Toxicology, **13** (1989), 235–244.

[ 3 ] Dunnett,C.W.: *A multiple comparisons procedure for comparing several treatments with a control.* J.Amer.Statist.Ass., **50** (1955), 1096–1121.

[ 4 ] Feller,W.: *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, (1966).

[ 5 ] Marcus,R., Peritz,E. and Gabriel,K.R.: *On closed testing procedures with special reference to ordered analysis of variances.* Biometrika, **63** (1976), 655–660.

[ 6 ] Marcus,R.: *The powers of some tests of the equality of normal means against an ordered alternative.* Biometrika, **63** (1976), 177–183.

[ 7 ] Plackett,R.L.: *The analysis of categorical data.* Griffin, London, second edition, (1981).

[ 8 ] Robertson,T., Wright,F.T. and Dykstra,R.T.: *Order Restricted Statistical Inference.* JOHN WILEY & SONS, (1987).

[ 9 ] Scheffé,H.: *A method for judging all contrasts in the analysis of variance.* Biometrika, **40** (1953), 87–104.

[10] Williams,D.A.: *A test for differences between treatment means when several dose levels are compared with a zero dose control.* Biometrics, **27** (1971), 103–117.

[11] Williams,D.A.: *Some inference procedures for monotonically ordered normal means.* Biometrika, **64** (1977), 9–14.

*Received November 25, 1993*
*Revised February 23, 1994*