Collapsing ordered categories for detecting the emergence of toxic chemical in fresh water by using dose-response curves

Yanagawa, Takashi
Kurume University, Biostatistics Center
67 Asahi-Machi
Kurume-Shi 830-0011,Japan
E-mail: yanagawa_takashi@kurume-u.ac.jp

Aoyama, Yoshiko
Kurume University, DGraduate School of Medicine
67 Asahi-Machi
Kurume-Shi 830-0011,Japan
E-mail: yoshiko.aoyama@gmail.com

SUMMARY

Trend tests in proportion are developed under binomial and extra-binomial variability for linear and non-linear dose-responses. Those tests employ orthonormal dose vectors and are proved to be score test, or generalized score test. They are applied for designing biological monitoring systems using water snails to detect the emergence of a toxic chemical in fresh water. Specifically, they are used for selecting the $2 \times K$ table from $a \times K$ tables ($a \geq 3$) that shows the strongest dose-response relationship among other $2 \times K$ tables.

Keywords binomial distribution; Cochran-Armitage test; generalized score test; Gram-Schimdt orthonormalization; orthonormal dose score; over dispersion.

1 Introduction

The present paper is concerned with an experiment for designing biological monitoring systems using water snails to detect the emergence of a toxic chemical in fresh water. Figure 1 shows an image of the monitoring system envisioned. The experiment is briefly described as follows. Adult freshwater snails (mean [± SD] of total length, 31.6 ± 1.3 mm) that were collected from Inose stream, Fukuoka, Japan, were kept under a photoperiod (12:12-h light:dark) for six months, and fed with artificial diet once a day were used in the experiment. They were exposed to copper sulfate (CuSO4; > 98% purity; 0, 0.1, 1 and 10 mg/L) in each chamber for 6 days that contained each level of test solution. The test solution in test chamber was changed once a day. See Kang et al. (2009) for details of the experiment.

Figure 2 shows five typical behavioral patterns of snails in water, from the stage of movement to the stage of withdrawal of the body into the shell, showing the levels from 'feeling well' to 'perceived risks'; more precisely, movement (M), adhesion and cessation of muscular activity (AC), detachment and muscular activity (DM), detachment and cessation of muscular activity (DC), withdrawal of the body into the shell (W). We consider those stages as ordered categories in this paper and call them simply as M, AC, DM, DC and W. The question asked is which behavior is the sharp indicator of the emergence of CuSO4, a well known toxic chemical.

Table 1 summarizes the data by doses and behaviors obtained in the experiment. The question may be statistically formulated as follows. As there are four possible classifications of categories into two, namely, (M1): {M} vs. {AC, DM, DC, W}, (M2): {M,AC} vs {DM, DC, W}, (M3): {M, AC, DM} vs {DC, W}, and (M4): {M, AC, DM, DC} vs. {W}, which one among those four is the most sensitive for detecting the emergence of the chemical compounds.

Now, if we select a cut-off point in lower categories, it is often the case that resulting $2 \times k$ tables show increasing trends of dose-response curves, but downturns are observed at higher doses when the cut-off point is selected in higher categories. Furthermore, data are potentially subject to over-dispersion since there might exist substantial day to day variation.

Cochran-Armitage (C-A) test (Cochran, 1954: Armitage, 1955) has been applied for detecting the trend of dose-response in proportions. However, if the data shows increasing fashion over low doses, but downturn in higher doses, the C-A test could lose powers for detecting the dose-response, as was pointed up by Simpson and Margolin (1986). Also it loses validity when over dispersion exists.

It is the purpose of the present paper to develop a method of collapsing categories in $a \times K$ table ($a \geq 3$) to get $2 \times K$ table that shows the strongest dose-response relationship among other collapses. We take different approach from Simpson and Margolin (1986). The problem is mathematically formulated in Section 2, score and generalized score tests (Boos, 1992) for linear and non-linear trends under potential existence of over-dispersion will be developed in the same section. Finally, the method will be applied to the data in Table 1.

2 Trend test

2.1 Assumptions and the problem

Consider a dose response experiment, where n_{ij} is the number of subjects (snails) that are administered dose d_i at jth day, $i = 1, 2, \ldots, k; j = 1, 2, \ldots, m_i$. Denote by Y_{ij}/n_{ij} the proportion of specified response. Assume that $\{Y_{i1}, Y_{i2}, \ldots, Y_{im_i}\}$ and $\{Y_{i'1}, Y_{i'2}, \ldots, Y_{i'm_{i'}}\}$ are independent for $i \neq i'$ and $d_1 < d_2 < \cdots < d_k$. Let π_i represents the response probability at d_i. Our goal is to test

$$H_0 : \pi_1 = \pi_2 = \cdots = \pi_k$$

against H_1 that postulates a linear or nonlinear trend in π's. Mathematical representation of H_1 will be given below.

2.2 Orthonormal dose vector

Let a dot in subscript denote the summation over that subscript, hence $n_{i\cdot} = \sum_j^{m_i} n_{ij}, Y_{i\cdot} = \sum_j Y_{ij}$ and so on. For i-th dose d_i, put

$$c_i = d_i - \bar{d},$$

where

$$\bar{d} = \sum_i d_i n_{i\cdot}/n_{\cdot\cdot},$$

and consider vector

$$\mathbf{c_1} = (c_1, c_2, \cdots, c_k).$$

Define

$$\mathbf{c_s} = (c_{s1}, c_{s2}, \cdots, c_{sk})',$$

where $c_{si} = c_i^s$ (sth power of c_i) for $s = 1, 2, \cdots, r$. Also define inner product of two vector by

$$(\mathbf{a}, \mathbf{b}) = \sum_i a_i b_i n_i$$

and its norm by

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})}.$$

It is obvious that $(\mathbf{c_0}, \mathbf{c_1}) = \mathbf{0}$ and that $\mathbf{c_0}, \mathbf{c_1}, \cdots, \mathbf{c_r}$ are linearly independent.

Let $\mathbf{a_0}, \mathbf{a_1}, \cdots, \mathbf{a_r}$ be orthonormal vectors obtained by applying Gram-Schimdt orthonormalization to ($\mathbf{c_0}$, $\mathbf{c_1}$, ..., $\mathbf{c_r}$) with respect to the innner product, that is

$$\mathbf{a_0} = \frac{\mathbf{c_0}}{\|\mathbf{c_0}\|}$$

$$\mathbf{d_s^*} = \mathbf{c_s} - \sum_{h=0}^{s-1} (\mathbf{c_s}, \mathbf{a_h})\mathbf{a_h}$$

$$\mathbf{a_s} = \frac{\mathbf{d_s^*}}{\|\mathbf{d_s^*}\|}$$

so we have

$$(\mathbf{a_s}, \mathbf{a_l}) = \begin{cases} 0 & \text{if } s \neq l \\ 1 & \text{if } s = l \end{cases}$$

and $\|\mathbf{a_s}\| = 1$ for all $s = 0, 1, \cdots, r$. We call $\mathbf{a_0}, \mathbf{a_1}, \cdots, \mathbf{a_r}$ the orthonormal dose score vectors.

2.3 The null and alternative hypotheses

Using orthonormal dose score vectors $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})'$, we introduce a logistic model by

$$\log \frac{\pi_i}{1 - \pi_i} = \sum_{s=0}^{r} \beta_s a_{si},$$

where β_s, $s = 1, 2, \ldots, r$, are unknown parameters. Putting $\boldsymbol{\beta}_{(2)} = (\beta_1, \cdots, \beta_r)'$, the null and alternative hypotheses for linear and non linear trend may be represented by

H$_0$: $\boldsymbol{\beta}_{(2)} = \mathbf{0}$ and H$_1$: $\boldsymbol{\beta}_{(2)} \neq \mathbf{0}$.

2.4 Trend Test under Binomial Variability

Suppose that Y_{ij} follows binomial distribution $B(n_{ij}, \pi_i), i = 1, \cdots, k$. For testing H_0 vs H_1 we propose a trend tests based on statistics

$$T_{Sr} = \sum_{s=1}^{r} \frac{(\sum_i a_{si} Y_{i.})^2}{\bar{Y}(1-\bar{Y})}, \qquad r = 1, 2, \cdots$$

where $Y_{i.} = \sum_{j=1}^{m_i} Y_{ij}$, $Y_{..} = \sum_{i=1}^{k} Y_{i.}$, $n_{..} = \sum_{i=1}^{k} \sum_{j=1}^{m_i} n_{ij}$, $\bar{Y} = Y_{..}/n_{..}$ and a_{si} is the ith element of the orthonormal dose score vector $\mathbf{a}_s$. It may be shown similarly as Jayasekara et al. (1999) that under H_0 the summands of T_{Sr} are mutually uncorrelated and that T_{Sr} follows a chi-square distribution with r degree of freedom, asymptotically. When $r = 1$, T_{Sr} is written as

$$S1 = \frac{(\sum_i a_{1i} Y_{i.})^2}{\bar{Y}(1-\bar{Y})}$$

which is easily shown equivalent to the statistics of the C-A test. When $r = 2$ T_{Sr} is written as

$$S2 = \sum_{s=1}^{2} \frac{(\sum_i a_{si} Y_{i.})^2}{\bar{Y}(1-\bar{Y})}.$$

We call the test based on this statistics the $S2$ test. We may prove the following proposition.

Proposition 1 T_{Sr} is the statistic of the score test for testing H_0 against H_1. T_{Sr} follows a chi-square distribution with r degree of freedom under H_0, asymptotically.

Proof The log likelihood function is represented by,

$$\ell(\boldsymbol{\beta}) = c + \sum_i \sum_j y_{ij} (\sum_{s=0}^{r} \beta_s a_{si}) - \sum_i \sum_j n_{ij} \log(1 - (1 + \exp\{-\sum_{s=0}^{r} \beta_s a_{si}\})^{-1})$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_r)$ and c is a constant term. Thus the proof of the proposition is immediate.

2.5 Trend test under over-dispersion

Boos(1992) introduced the generalization of score test that able to account for certain model inadequacies or lack of knowledge by use of empirical variance estimates. We apply his idea to obtain trend test when data are subject to over-dispersion. We have the following theorem.

Proposition 2 For orthonormal dose vector $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})', s = 1, 2, \cdots, r$, put

$$S = \left(\sum_i^K a_{1i} Y_{i.}, \cdots, \sum_i^K a_{ri} Y_{i.} \right)'$$

and

$$D = \left(\sum_i^K \sum_{j=1}^{m_i} (Y_{ij} - n_{ij}\bar{Y})^2 a_{ti} a_{ui} \right)_{r \times r}.$$

Then

$$T_{GSr} \;=\; S' D^{-1} S$$

is the generalized score test for testing H$_0$: $\boldsymbol{\beta}_{(2)} = \mathbf{0}$ vs. H$_1$: $\boldsymbol{\beta}_{(2)} \neq \mathbf{0}$. T_{GSr} follows chi-square distribution with r degree of freedom under H_0, asymptotically.

(The proof of the theorem is given in Appendix.)

When $r = 1$, T_{GSr} is written as

$$GS1 = \frac{(\sum_i^K a_{1i} Y_{i\cdot})^2}{\sum_i^K a_{1i}^2 \sum_{j=1}^{m_i} (Y_{ij} - n_{ij}\bar{Y})^2}$$

which is easily shown equivalent to the generalized C-A test (Carr and Gorelick , 1995), thus has high power in detecting linear trend. We call the test based on GS1 the gC-A (generalized C-A) test.

When $r = 2$ T_{GSr} is written as

$$GS2 \quad = \quad \left(\sum_i^K a_{1i} Y_{i\cdot}, \quad \sum_i^K a_{2i} Y_{i\cdot} \right) \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}^{-1} \begin{pmatrix} \sum_i a_{1i} Y_{i\cdot} \\ \sum_i a_{2i} Y_{i\cdot} \end{pmatrix}$$

where

$$v_{sl} = \sum_i^K a_{si} a_{li} \sum_{j=1}^{m_i} (Y_{ij} - n_{ij}\bar{Y})^2$$

We call the test based on this statistics the $GS2$ test. We propose it for testing $H_0 : \pi_1 = \cdots = \pi_k$ against a convex trend. Note that no specific distribution is assumed for the generalized score test. However, in below we examine the behaviour of the GS1 and GS2 tests under beta-binomial distribution, assuming that $\{Y_{ij}\}_{j=1,\cdots,m_i}$ are independent and Y_{ij} follows,

$$P(Y_{ij} = y) = \begin{pmatrix} n_{ij} \\ y \end{pmatrix} \frac{\Gamma(y + (\pi_i/\varphi))\Gamma(n_{ij} - y + (1 - \pi_i)/\varphi)\Gamma(1/\varphi)}{\Gamma(\pi_i/\varphi)\Gamma((1 - \pi_i)/\varphi)\Gamma((1/\varphi) + n_{ij})},$$

where φ is the dispersion parameter.

3 Numerical Evaluation

We consider linear or quadratic responses shown in Table 1. We call the response patterns of No. 1, 2, 3 and 4 in Table 2, the uniform, convex, concave and increasing monotone response, respectively.

Performance of the C-A test, S2 tests gC-A test GS2 tests are examined in terms of the Type I error and powers for detecting the true patterns by simulation. We consider two beta-binomial distributions with response probabilities given in Table 2 and with the value of dispersion parameter φ 0.1 and 1, respectively. The significance level is taken as 0.05. 10,000 data are generated from each distribution, and empirical Type I errors and the powers are computed when $n_{ij} = 10, i, j = 1, 2, \ldots, 5$. Table 3 summarizes the results. The table indicates the following results;

1. When an over-dispersion exists, the sizes of C-A and S2 tests substantially inflate. The powers of those tests are remarkably high, but those powers loose the validity since the inflation of the sizes of tests.

2. On the other hand sizes of gC-A and GS2 tests are kept near the nominal level (5% in the table) whether an over-dispersion exists or not.

3. The gC-A has higher power than the GS2 test for monotone increasing trends, but has lower power than the GS2 test for convex alternatives that include down-turns at high doses.

4 Application

Values of C-A, $S2$, gC-A and $GS2$, and also their p-values, computed from Table 1 are listed in Table 4. The table shows that all p-values in the table except gC-A at M4 are less than 5%, confirming that CuSO4 is a toxic chemical compounds. It also shows that p-values of C-A and S2 tests are extremely

small, indicating the existence of over-dispersion, and p-values of these test should not be taken into account seriously. p-values of gC-A at M3 and at M4 are almost equal, p-value of GS2 at M2 and at M3 are almost equal, and it is not easy to select M2 and M3, but we selected M3 based on the p-value of GS2, that is 0.0033. It is suggested to classify four categories into {M, AC, DM} vs. {DC, W} and employ GS2 test for detecting the emergence of the chemical in the biological monitoring system.

Appendix

Proof of Proposition 2

Let

$$\begin{aligned} S &= S(\boldsymbol{\beta}) = \left(\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_t}\right)_{(r+1)\times 1} \\ I_Y &= I_Y(\boldsymbol{\beta}) = \left(-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_t \partial \beta_u}\right)_{(r+1)\times(r+1)} \\ D_Y &= D_Y(\boldsymbol{\beta}) = \left(\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_t}\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_u}\right)_{(r+1)\times(r+1)} \end{aligned}$$

Reperent $S' = (S'_{(1)}, S'_{(2)})$, where $S_{(1)}$ is 1×1 and $S_{(2)}$ is $r \times r$. The matrices above are partitioned accordingly, e.g., $I_{Y(11)}$ is 1×1, $I_{Y(12)}$ is $1 \times r$ and so on. For testing $H_0 : \boldsymbol{\beta}_{(2)} = \mathbf{0}$ vs $H_1 : \boldsymbol{\beta}_{(2)} \neq \mathbf{0}$, Boos (1992) proposed generalized score test as:

$$T_{GS} = \tilde{S}'_{(2)} \tilde{\mathbf{V}}(\tilde{S}_{(2)}) \tilde{S}_{(2)} \qquad (A1)$$

where $\tilde{\cdot}$ denote those matrices evaluated at $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$, and $\tilde{\boldsymbol{\beta}}$ is restricted mle of $\boldsymbol{\beta}$ under H_0, and

$$\begin{aligned} \tilde{\mathbf{V}}(\tilde{S}_{(2)}) &= \tilde{D}_{Y(22)} - \tilde{I}_{Y(21)}\tilde{I}^{-1}_{Y(11)}\tilde{D}'_{Y(21)} - \tilde{D}_{Y(21)}\tilde{I}^{-1}_{Y(11)}\tilde{I}'_{Y(21)} \\ &\quad + \tilde{I}_{Y(21)}\tilde{I}^{-1}_{Y(11)}\tilde{D}_{Y(11)}\tilde{I}^{-1}_{Y(11)}\tilde{I}'_{Y(21)} \end{aligned}$$

Since $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})', s = 1, 2, \cdots, r$ are orthonormal, it follows that $\tilde{I}_{Y(21)} = \tilde{I}'_{Y(12)} = \mathbf{0}$. Thus we have $\tilde{\mathbf{V}}(\tilde{S}_{(2)}) = \tilde{D}_{Y(22)}$. Furthermore, a straightforward computation shows that

$$\tilde{S}_{(2)} = \left(\sum_i a_{1i} Y_{i\cdot}, \cdots, \sum_i a_{ri} Y_{i\cdot}\right)'$$

Thus submitting those equation to (A1) we have the desired result.

REFERENCES

Armitage, P. (1955) Test for linear trend in proportions and frequencies *Biometrics, 11, 375-386.*

Boos, D.D. (1992) On generalized score test *The American Statistician, 46, 327-333.*

Carr, G.J and Gorelick, N.J. (1995) Statistical design and analysis of mutation studies in transgenic mice *Environmental and Molecular Mutagenesis, 25, 245-255.*

Cochran, W.G. (1992) Some methods for strengthning the common χ^2 tests *Biometrics, 10, 417-451.*

Jayasekara, L.,Yanagawa, T. and Tsujitani, M. (1994) A Location-dispersion test for $2 \times k$ tables *Bulletin of Informatics and Cybernetics, 26,125-139.*

Kang, J. Ik, Nakamura, A., Moroishi, J., Ishibashi, K., Fukuda, S., Shimasaki , Y. and Oshima, Y. (2009) Effects of heavy metal compounds on behavior of freshwater snail (*Semisulcospira libertine) (in Japanese) Scientific Bulletin, Faculty of Agriculture, Kyushu University, 64-2, 119-123.*

Margolin, B.H., Kaplan, N. andZeiger, E. (1981) Statistical analysis of the Ames Salmonera/microsome test. Proceedings of the National Academy of Sciences, USA, 76, 3779-3783.

Simpson, D. G. and Margolin, B. H. (1986): Recursive nonparametric testing for dose-response relationship subject to down turns at high doses; Biometrika, 73,586-596.

Figures and Tables

Figure 1: Biological monitoring system

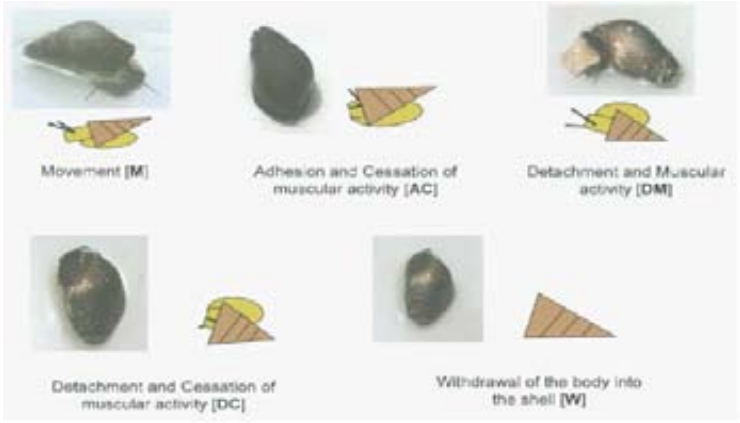

Figure 2: Five stages of snails

Table 1: Behavior of snails in the experiment

day	dose	M	AC	DM	DC	W	Total
1	0	7	0	0	0	1	8
	0.1	4	3	0	0	1	8
	1	0	0	0	0	8	8
	10	0	0	0	0	8	8
2	0	2	2	1	1	2	8
	0.1	0	0	0	0	8	8
	1	0	0	0	0	8	8
	10	0	0	0	0	8	8
3	0	5	1	0	1	1	8
	0.1	0	3	0	1	4	8
	1	0	0	0	0	8	8
	10	0	0	0	2	6	8
4	0	6	0	0	0	2	8
	0.1	0	0	0	0	8	8
	1	1	0	0	0	7	8
	10	0	0	0	3	5	8
5	0	5	2	0	0	1	8
	0.1	0	0	0	0	8	8
	1	0	0	0	2	6	8
	10	0	0	0	2	6	8
6	0	6	0	0	0	2	8
	0.1	0	0	0	0	8	8
	1	0	0	0	1	7	8
	10	0	0	0	1	7	8

No.	π_1	π_2	π_3	π_4	π_5	Response Patterns
1	0.2	0.2	0.2	0.2	0.2	uniform
2	0.2	0.2	0.27	0.4	0.6	increasing monotone
3	0.2	0.45	0.8	0.6	0.5	convex
4	0.8	0.6	0.5	0.5	0.6	concave

Table 2: Trend patterns of population probabilities

Binomial	C-A	S2	gC-A	GS2
Type I error	0.051	0.051	0.050	0.037
monotone increasing	0.997	0.997	0.989	0.954
convex	0.938	1.000	0.682	0.992
concave	0.634	0.919	0.452	0.747
Beta-Binomial ($\phi = 0.1$)	C-A	S2	gC-A	GS2
Type I error	0.143	0.185	0.048	0.036
monotone increasing	0.983	0.985	0.891	0.767
convex	0.873	0.999	0.499	0.908
concave	0.607	0.877	0.316	0.503
Beta-Binomial ($\phi = 1$)	C-A	S2	gC-A	GS2
Type I error	0.406	0.599	0.040	0.029
monotone increasing	0.891	0.934	0.484	0.350
convex	0.758	0.971	0.243	0.522
concave	0.600	0.847	0.150	0.239

Table 3: Trend patterns, Type I errors and powers

	C-A	$S2$	gC-A	$GS2$
M1	18.90	46.79	6.06	8.75
	(1.4E-05)	(6.9E-11)	(0.014)	(0.013)
M2	26.47	64.48	6.97	11.23
	(2.7E-07)	(1.0E-14)	(0.0083)	(0.0037)
M3	27.24	66.57	7.04	11.43
	(1.8E-07)	(3.6E-15)	(0.0078)	(0.0033)
M4	10.26	41.66	3.75	8.89
	(0.001)	(9.0E-10)	(0.053)	(0.012)

Table 4: Values of statistics and their p-values in the blankets

M1; {M} vs. {AC, DM, DC, W}, M2; {M,AC} vs {DM, DC, W},
M3; {M, AC, DM} vs {DC, W}, M4; {M, AC, DM, DC} vs. {W}