应用概率统计　第二十五卷
第五期　2009年10月

Chinese Journal of Applied Probability
and Statistics Vol.25 No.5 Oct. 2009

Assessing Occupational Exposure Via the Unbalanced One-Way Random Models *

Li Xinmin

(*Institute of Applied Mathematics, School of Science, Shandong University of Technology, Zibo, 255049*)

Abstract

A unbalanced one-way random model is considered for assessing the proportion of workers whose mean exposure exceed the occupational exposure limit (OEL) based on exposure measurements to the worker. Hypothesis testing for the relevant parameter of interest is proposed when the exposure data are unbalanced. The method is based on the generalize inference. A simulation study is conducted to compare it with that of Krishnamoorthy and Guo (2005). Simulation results suggest that the proposed method appears to be better, especially in very unbalanced design.

Keywords: Generalized p-value, power, random effect, occupational exposure limit.

AMS Subject Classification: 62F03, 62F25.

§1. Introduction

Exposure levels at workplaces are commonly assessed by the proportion of exposure measurements exceeding the occupational exposure limit (OEL). Some of recent work on exposure monitoring has focused on the use of the one-way variance component model to incorporate the between works and within works source of variability (see Rapport, et al. (1993), Rapport, et al. (1995), Lyles, et al. (1997a, b) and Maxim, et al. (2000)). When the model is used to analyze logged exposure data, the interested parameter (the proportion of measurements exceeding a OEL) is a function of overall mean and the variance components in the model. A hypothesis testing problem involving this parameter is addressed in Lyles, et al. (1997a, b), they proposed some large sample approximation methods.

The concepts of generalized inference introduced by Tsui and Weerahandi (1989) and Weerahandi (1993) appear to be appropriate for above hypothesis testing problem. Recently, Krishnamoorthy and Mathew (2002) used the generalized inference approach

*This work was partly supported by National Natural Science Foundation of China (10771126) and Shandong University of Technology (2006KJZ01).

Received April 7, 2006. Revised June 5, 2008.

for testing problem with balanced data. Mu, et al. (2007) considers a balanced two-way random effects model for studying the mean exposure level of a factory. As pointed out by Lyles, et al. (1997b), collecting balanced data seems to be unrealistic because of the workers burden, and unavailability of the workers during the sampling period. Furthermore, exposure data sets are typically small, and hence methods are really needed for small unbalanced data. More recently, Krishnamoorthy and Guo (2005) provided a generalized p-value approach for testing problem applying the results of Thomas and Hultquist (1978) when the data are unbalanced. However, the results of Thomas and Hultquist (1978) suggested that the approximation is not work well in cases where the variance components ratio (of the between variance and the within variance) less than 0.25 and the data is extremely unbalanced.

In this paper, we provide a different approach to construct the generalized p-value using the results of Li and Li (2005). The article is organized as follows. A brief review of the generalized inference is given in next section. Furthermore, the unbalanced one-way random model and the problem of hypothesis testing of interested parameter are introduced. Section 3 presents the hypothesis testing problem along with a solution based on a generalized test variable for the proportion of exposure measurements. In Section 4, a simulation study is employed to compare the proposed method given in Section 3 with the method by Krishnamoorthy and Guo (2005). The sizes and powers are numerically evaluated and presented. Simulation results indicate that the proposed method appears to be better, especially in the cases where the variance components ratio less than 0.25 and the data is extremely unbalance.

§2. Preliminaries

2.1 Generalized p-value

Tsui and Weerahandi (1989) introduced the concept of generalized inference for testing hypothesis. Consider an observable random vector X with a probability distribution $\mathsf{P}_\eta(\cdot)$, where $\eta = (\theta, \delta)$ is an unknown vector in parameter space Ω, $\theta = \theta(\eta)$ is a real-valued parameter of interest, and δ is the nuisance parameter. Assume that Θ is the parameter space of θ, and x is the observed value of X. The problem of interest is to test the one-sided hypothesis $H_0 : \theta \le \theta_0$ versus $H_1 : \theta > \theta_0$.

Definition 2.1 Let $R = R(X; x, \eta)$ be a function of X, x, and $\eta = (\theta, \delta)$. R is said to be a generalized test variable if it has the following properties:

(a) The observed value $r = R(x; x, \eta)$ does not depend on the nuisance parameter δ.

(b) R has a probability distribution free of unknown parameters.

(c) For fixed x and δ, $\mathsf{P}\{R(X; x, \eta) \geq r|\theta\}$ is nondecreasing in θ.

Following condition Property (a)–(c) of Definition 2.1, a 'large' observed value $r = R(x; x, \eta)$ suggests evidence against H_0, we can used the generalized p-value

$$p = \sup_{\theta \leq \theta_0} \mathsf{P}\{R(X; x, \eta) \geq r|\theta\} = \mathsf{P}\{R(X; x, \eta) \geq r|\theta_0\}$$

to test $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$, and small p-value indicates that the observed does not support H_0.

2.2 The Model for Exposure Assessment

We now describe the one-way variance components model to the exposure data given in Rapport (1995) and Lyles, et al. (1997a). Let X_{ij} denote the jth shift-long exposure measurement for the ith worker, assumed to be distributed as lognormal distribution, $j = 1, 2, \cdots, b_i$, $i = 1, 2, \cdots, a$. Let $Y_{ij} = \ln(X_{ij})$, so that Y_{ij} follows a normal distribution. Then the one-way random effect model for the Y_{ij} is given as

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \qquad j = 1, 2, \cdots, b_i,\ i = 1, 2, \cdots, a, \tag{2.1}$$

where μ is the overall mean, α_i's and ε_{ij}'s are mutually independent normal random variables with zero means and variances σ_α^2 and σ_ε^2, respectively. Define $Y = (Y_{11}, \cdots, Y_{ab_a})'$, $Z = \operatorname{diag}(1_{b_1}, \cdots, 1_{b_a})$, and $n = \sum\limits_{i=1}^{a} b_i$, the model (2.1) can be written as

$$Y = \mu 1_n + Z\alpha + \varepsilon, \tag{2.2}$$

where $1_a \in R^a$ is a matrix with all elements being 1. Thus $Y \sim N(\mu 1_n, \sigma_\alpha^2 Z'Z + \sigma_\varepsilon^2 I_n)$.

Let α_i represents the random effects due to the ith worker, the mean exposure for the ith worker is given as follows

$$\mu_{x_i} = \mathsf{E}(X_{ij}|\alpha_i) = \mathsf{E}(\exp(Y_{ij})|\alpha_i) = \exp(\mu + \alpha_i + \sigma_\varepsilon^2/2). \tag{2.3}$$

Let θ be the probability that mean exposures exceed the occupational limit (OEL). Then, we have

$$\theta = \mathsf{P}(\mu_{x_i} > \mathrm{OEL}) = \mathsf{P}(\ln(\mu_{x_i}) > \ln(\mathrm{OEL})) = 1 - \Phi\Big(\frac{\ln(\mathrm{OEL}) - \mu - \sigma_\varepsilon^2/2}{\sigma_\alpha}\Big), \tag{2.4}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. The hypothesis of interest in our problem is

$$H_0: \theta \geq A \qquad \text{vs.} \qquad H_1: \theta < A,$$

where A is a specified quantity, usually small. It follows from (2.4) that the above hypothesis is equivalent to

$$H_0: \eta \geq \ln(\text{OEL}) \qquad \text{vs.} \qquad H_1: \eta < \ln(\text{OEL}), \tag{2.5}$$

where $\eta = \mu + z_{1-A}\sigma_\alpha + \sigma_\varepsilon^2/2$, and z_{1-A} is the $100(1-A)$th percentile of the standard normal distribution.

§3. Proposed Method

We shall now derive a generalized test variable for the hypothesis testing (2.5). Define $\overline{Y}_{i\cdot} = \sum\limits_{j=1}^{b_i} Y_{ij}/b_i$, $\overline{Y}_{\cdot\cdot} = \sum\limits_{i=1}^{a}\sum\limits_{j=1}^{b_i} Y_{ij}/n$, $S_1 = \sum\limits_{i=1}^{a} b_i(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})^2$, and $S_2 = \sum\limits_{i=1}^{a}\sum\limits_{j=1}^{b_i}(Y_{ij} - \overline{Y}_{i\cdot})^2$. When model (2.1) is balanced, that is, $b_1 = b_2 = \cdots = b_a = b$, S_1 and S_2 have independent scaled chi-squared distributions. Inference on the variance components is based on these distributional properties, see Krishnamoorthy and Mathew (2002). In the unbalanced case, S_1 and S_2 are still independent, and S_2 still has a scaled chi-squared distribution. However, unless $\sigma_\alpha^2 = 0$, S_1 no longer has a scaled chi-squared distribution.

Thomas and Hultquist (1978) recommended a statistic S_{1u}, instead of S_1, where

$$S_{1u} = b_h \sum_{i=1}^{a}(\overline{Y}_{i\cdot} - \overline{Y}^*_{\cdot\cdot})^2, \qquad \overline{Y}^*_{\cdot\cdot} = \sum_{i=1}^{a}\overline{Y}_{i\cdot}/a, \qquad b_h = a\Big/\Big(\sum_{i=1}^{a} 1/b_i\Big). \tag{3.1}$$

Set $\theta_{1u} = \mathsf{E}(S_{1u}/n_1) = \sigma_\varepsilon^2 + b_h\sigma_\alpha^2$. The term S_{1u} represents the unweighed sum of squares of the treat means and b_h represents the harmonic mean of the b_i values. They showed that S_{1u}/θ_{1u} is well approximated by χ^2_{a-1} except in cases where the variance components ratio $\lambda = \sigma_\alpha^2/\sigma_\varepsilon^2 < 0.25$ and the design is extremely unbalanced.

Using this fact, Krishnamoorthy and Guo (2005) proposed the generalized p-value for hypothesis testing (2.5) using the approximation generalized test variables. Unfortunately, using the above results, we see that the test is not satisfactory in cases where $\lambda = \sigma_\alpha^2/\sigma_\varepsilon^2 < 0.25$ and the design is extremely unbalanced. Now we apply the results of Li and Li (2005) to reconsider the problem.

Set $Q_1 = (Z'Z)^{-1}Z' = \text{diag}((1/b_1)1'_{b_1}, \cdots, (1/b_a)1'_{b_a})$ be an $a \times n$ matrix, and Q_2 be an $(n-a) \times n$ matrix such that $Q_1Q'_2 = 0$, and $Q_2Q'_2 = I_{n-a}$. It follows that

$$\begin{aligned}&\overline{Y} = Q_1Y \sim N(\mu 1_a, \sigma_\alpha^2 I_a + \sigma_\varepsilon^2 D),\\&Q_2Y \sim N(0, \sigma_\varepsilon^2 I_{n-a}),\ \overline{Y} \text{ and } Q_2Y \text{ are mutually independent,}\end{aligned}$$

where $D = \text{diag}(1/b_1, 1/b_2, \cdots, 1/b_a)$, and I_m is an identify matrix of order m.

Let $H = (H'_1, H'_2)'$ with $H_1 = (1/\sqrt{a}) \cdot 1'_a$ and H_2 be an $(a-1) \times a$ matrix such that $H_1H'_2 = 0$, and $H_2H'_2 = I_{a-1}$. Then we have

$$\begin{pmatrix} H_1\overline{Y} \\ H_2\overline{Y} \end{pmatrix} \sim N\left(\begin{pmatrix} \sqrt{a}\mu \\ 0 \end{pmatrix}, \sigma_\alpha^2 I_a + \sigma_\varepsilon^2 HDH'\right). \tag{3.2}$$

Set $T = H_2\overline{Y}$, and $\Lambda = H_2DH'_2$. It is easy to see that $\overline{Y}^*_{\cdot\cdot} = (1/\sqrt{a}) \cdot H_1\overline{Y}$, $S_{1u} = \overline{Y}'H'_2H_2\overline{Y} = T'T$, and $S_2 = Y'Q'_2Q_2Y$.

Following from (3.2), we have

$$\begin{pmatrix} \overline{Y}^*_{\cdot\cdot} \\ \\ T \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \\ 0 \end{pmatrix}, \begin{pmatrix} \dfrac{1}{ab_h}(b_h\sigma_\alpha^2 + \sigma_\varepsilon^2) & \dfrac{1}{\sqrt{a}}\sigma_\varepsilon^2 H_1DH'_2 \\ \dfrac{1}{\sqrt{a}}\sigma_\varepsilon^2 H_2DH'_1 & \sigma_\alpha^2 I_{a-1} + \sigma_\varepsilon^2\Lambda \end{pmatrix}\right), \quad S_2 \sim \sigma_\varepsilon^2\chi^2_{n-a}, \tag{3.3}$$

where $(\overline{Y}^*_{\cdot\cdot}, T)$ and S_2 are mutually independent. It is easy to see $(\overline{Y}^*_{\cdot\cdot}, T, S_2)$ is a sufficient statistic. Hence the interval estimation can be constructed based on them. Let $\overline{y}^*_{\cdot\cdot}$, t and s_2 denote the observed values of the random $\overline{Y}^*_{\cdot\cdot}$, T and S_2. Following (3.3), define

$$\begin{aligned}R &= r(Y; y, \eta)\\&= \overline{y}^*_{\cdot\cdot} - \frac{\overline{Y}^*_{\cdot\cdot} - \mu}{\sqrt{b_h\sigma_\alpha^2 + \sigma_\varepsilon^2}}\Big[b_h\frac{s_{1u} - T'(\sigma_\alpha^2\Lambda^{-1} + \sigma_\varepsilon^2 I_{a-1})^{-1}T \cdot s_2\sigma_\varepsilon^2/S_2}{T'(\sigma_\alpha^2 I_{a-1} + \sigma_\varepsilon^2\Lambda)^{-1}T} + \frac{s_2\sigma_\varepsilon^2}{S_2}\Big]_+^{1/2}\\&\quad + z_p\Big[\frac{s_{1u} - T'(\sigma_\alpha^2\Lambda^{-1} + \sigma_\varepsilon^2 I_{a-1})^{-1}T \cdot s_2\sigma_\varepsilon^2/S_2}{T'(\sigma_\alpha^2 I_{a-1} + \sigma_\varepsilon^2\Lambda)^{-1}T^{1/2}}\Big]_+^{1/2} + \frac{s_2\sigma_\varepsilon^2}{2S_2} - \eta,\end{aligned}$$

where $[a]_+ = \max\{a, 0\}$, $(\overline{Y}^*_{\cdot\cdot} - \mu)/\sqrt{b_h\sigma_\alpha^2 + \sigma_\varepsilon^2} \sim N(0,1)$ and $(\sigma_\alpha^2 I_{a-1} + \sigma_\varepsilon^2\Lambda)^{-1/2}T \sim N(0, I_{a-1})$ are dependent. Let $W_0 \sim N(0,1)$, $W_1 \sim N(0, I_{a-1})$, $W_2 \sim \chi^2_{n-a}$, and W_0, W_1 and W_2 are mutually independent. Then we have

$$\begin{aligned}R &\overset{\text{d}}{\sim} \overline{y}^*_{\cdot\cdot} - \frac{W_0}{\sqrt{ab_h}}\Big[b_h\frac{s_{1u} - W'_1\Lambda W_1 s_2/W_2}{W'_1W_1} + \frac{s_2}{W_2}\Big]_+^{1/2}\\&\quad + z_p\Big[\frac{s_{1u} - W'_1\Lambda W_1 s_2/W_2}{W'_1W_1}\Big]_+^{1/2} + \frac{s_2}{2W_2} - \eta\\&= R^* - \eta,\end{aligned} \tag{3.4}$$

where $\overset{\mathrm{d}}{\sim}$ denotes "approximately distributed".

Using the first expression in (3.4), it is easy to see that the observed value of R is 0, and the distribution of R is stochastically monotone in η. Although the second expression in (3.4) has a distribution free of any unknown parameters, but the actual distribution of R does depend on unknown parameters. However, using the second expression in (3.4), R is an approximate generalized test variable. Thus the generalized p-value for testing (2.5) is defined as

$$p = \mathsf{P}(R \geq 0|\eta = \ln(\mathrm{OEL})) = \mathsf{P}(R^* \leq \ln(\mathrm{OEL})|\eta = \ln(\mathrm{OEL})).$$

Since the distribution of R using second expression, given $\overline{y}^*_{\cdot\cdot}$, s_{1u}, s_2, is free of any unknown parameters, the Monte Carlo method can be used to compute the generalized p-value. The following algorithm can be used for obtaining it.

Algorithm 1:

(1) For a given data set, compute $\overline{y}^*_{\cdot\cdot}$, s_{1u}, s_2;

(2) For $j = 1, 2, \cdots, N$, generate

$$W_0 \sim N(0,1), \qquad W_1 \sim N(0, I_{a-1}), \qquad W_2 \sim \chi^2_{n-a};$$

(3) Compute the corresponding value R^*_j using the second expression in (3.4);

(4) End j loop.

Then the simulated generalized p-value for testing (2.5) is [Number of $R^*_j > \ln(\mathrm{OEL})$]/$N$.

Remark 1 If all the b_i's are equal, it can be easily verified that the R in (3.4) simplifies to the generalized pivot variable given in Krishnamoorthy and Mathew (2002).

§4. Simulation Study

We shall now study the size and power properties of the generalized test. The method (LI) described in Section 3 is now compared with that provided by Krishnamoorthy and Guo (2005) (KG) through a simulation study. The criteria for analyzing the performance of the methods are to compare the sizes and powers properties of tests.

In the simulation, five unbalanced patterns were selected, which are shown in Table 1 and 2. Without loss of generality, $\mu = 0$ is assigned. Note that

$$\begin{aligned}&\overline{Y}_{1\cdot}, \cdots, \overline{Y}_{a\cdot} \text{ i.i.d. } \sim N(0, (\lambda + 1/b_i)\sigma^2_\varepsilon),\\ &S_2 \sim \sigma^2_\varepsilon \chi^2_{n-a},\ (\overline{Y}_{1\cdot}, \cdots, \overline{Y}_{a\cdot}) \text{ and } S_2 \text{ are mutually independent,}\end{aligned} \tag{4.1}$$

where $\lambda = \sigma_\alpha^2/\sigma_\varepsilon^2$. Also, $\overline{Y}_{\cdot\cdot}^*$ and S_{1u} depend on only $\overline{Y}_{i\cdot}$'s, b_i's and a. For each unbalanced pattern and σ_α^2, σ_ε^2, and A, a Monte Carlo method of evaluating the size and power of the proposed test in Section 3 is given in the following algorithm.

Algorithm 2:

(1) For $i = 1$ to M, generate $(\overline{y}_{1\cdot}, \cdots, \overline{y}_{a\cdot}, s_2)$ according to (4.1), and compute $\overline{y}_{\cdot\cdot}^*$ and s_{1u} by (3.1), respectively;

(2) Use Algorithm 1 to compute the generalized p-value p_i;

(3) End i loop;

(4) If the parameters μ, σ_α^2, σ_ε^2, and A are chosen such that $\eta = \ln(\text{OEL})$, [Number of $p_i < \alpha$]/M is a simulated estimate of the size; if the parameters μ, σ_α^2, σ_ε^2, and A are chosen such that $\eta < \ln(\text{OEL})$, [Number of $p_i < \alpha$]/M is a simulated estimate of the power.

The Monte Carlo method of evaluating the size and power of the test can be found in Section 4 of Krishnamoorthy and Guo (2005).

Tables 1 and 2 give the size and the power values of the test at the significant level $\alpha = 0.05$ with $M = 5000$ and $N = 10000$. For computing the size, we choose $\theta = A = 0.05$. For computing the power, we choose $\theta = 0.002$ and $A = 0.05$ in Table 1, and choose $A = 0.05$ and $\mu_x/\text{OEL} = 0.2$ in Table 2, where the mean exposure $\mu_x = \exp(\mu + (\sigma_\alpha^2 + \sigma_\varepsilon^2)/2)$.

The numerical results in Table 2 show that both LI and KG method can have the sizes below the nominal level when λ is not too small. The sizes by LI test maintain stated level across all values of λ for all patterns, although it produces very conservative interval when λ is small. In contrast, the sizes by KG test may be exceeding the nominal level. Especially, the simulated sizes by KG sizes are very large in cases with small λ and very unbalanced designs. The powers of the two tests are increasing with the sample sizes. Although the powers by LI test are sometimes little than ones by KG test, they are generally vary little.

The simulation study indicates that both the LI and KG tests are useful for large values of λ, and the LI method is good for all values of λ. In summary, the LI test are recommended for extremely unbalanced designs in situations where λ is thought to be small. In any other situation, LI and KG tests can be recommended.

Acknowledge The author cordially thank referees for their valuable comments which lead to the improvement of this paper.

Table 1 Monte Carlo estimates of the sizes and powers

a	$(b_1,\cdots,b_a)$	$\sigma_\alpha^2/\sigma_\varepsilon^2$	σ_ε^2	Size LI	Size KG	Power LI	Power KG
3	(5,10,15)	0.05	0.5	0.037	0.039	0.077	0.078
		0.10	0.5	0.044	0.046	0.100	0.098
		0.25	0.5	0.051	0.049	0.115	0.111
		0.50	0.5	0.051	0.051	0.122	0.118
		0.75	0.5	0.047	0.052	0.136	0.131
		1.00	0.5	0.049	0.047	0.123	0.117
10	(1 2s,10 8s)	0.05	0.5	0.044	0.095	0.145	0.207
		0.10	0.5	0.055	0.094	0.202	0.248
		0.25	0.5	0.054	0.073	0.296	0.314
		0.50	0.5	0.055	0.065	0.381	0.388
		0.75	0.5	0.053	0.057	0.417	0.424
		1.00	0.5	0.046	0.049	0.433	0.433
10	(1 2s,4 2s,6 2s,8 2s,10 2s)	0.05	0.5	0.040	0.066	0.121	0.168
		0.10	0.5	0.054	0.080	0.171	0.210
		0.25	0.5	0.052	0.072	0.260	0.291
		0.50	0.5	0.051	0.063	0.328	0.351
		0.75	0.5	0.046	0.054	0.376	0.396
		1.00	0.5	0.049	0.054	0.403	0.424
16	(1 4s,2 4s,8 8s)	0.05	0.5	0.046	0.084	0.143	0.207
		0.10	0.5	0.058	0.090	0.205	0.255
		0.25	0.5	0.054	0.084	0.364	0.396
		0.50	0.5	0.055	0.065	0.490	0.501
		0.75	0.5	0.053	0.060	0.565	0.573
		1.00	0.5	0.052	0.055	0.616	0.621
20	(1 5s,4 5s,8 5s,12 5s)	0.05	0.5	0.051	0.103	0.155	0.228
		0.10	0.5	0.056	0.103	0.247	0.290
		0.25	0.5	0.050	0.080	0.431	0.450
		0.50	0.5	0.056	0.073	0.604	0.612
		0.75	0.5	0.051	0.062	0.678	0.680
		1.00	0.5	0.050	0.058	0.716	0.729

Table 2 Monte Carlo estimates of the sizes and powers

a	$(b_1,\cdots,b_a)$	$\sigma_\alpha^2/\sigma_\varepsilon^2$	σ_ε^2	Size		Power	
				LI	KG	LI	KG
3	(5,10,15)	0.05	0.5	0.033	0.035	0.577	0.552
		0.10	0.5	0.041	0.042	0.508	0.483
		0.25	0.5	0.047	0.046	0.350	0.331
		0.50	0.5	0.050	0.047	0.250	0.236
		0.75	0.5	0.050	0.053	0.209	0.196
		1.00	0.5	0.049	0.047	0.161	0.150
10	(1 2s,10 8s)	0.05	0.5	0.040	0.082	0.982	0.983
		0.10	0.5	0.053	0.090	0.969	0.970
		0.25	0.5	0.055	0.074	0.939	0.942
		0.50	0.5	0.054	0.066	0.834	0.836
		0.75	0.5	0.052	0.059	0.713	0.717
		1.00	0.5	0.046	0.049	0.602	0.604
10	(1 2s,4 2s,6 2s,8 2s,10 2s)	0.05	0.5	0.033	0.056	0.974	0.974
		0.10	0.5	0.046	0.070	0.967	0.967
		0.25	0.5	0.051	0.070	0.910	0.919
		0.50	0.5	0.051	0.062	0.778	0.790
		0.75	0.5	0.046	0.053	0.658	0.674
		1.00	0.5	0.048	0.052	0.552	0.571
16	(1 4s,4 4s,8 8s)	0.05	0.5	0.040	0.070	0.998	0.998
		0.10	0.5	0.054	0.083	0.996	0.997
		0.25	0.5	0.057	0.083	0.988	0.988
		0.50	0.5	0.055	0.066	0.956	0.957
		0.75	0.5	0.053	0.060	0.876	0.878
		1.00	0.5	0.052	0.055	0.799	0.803
20	(1 5s,4 5s,8 5s,12 5s)	0.05	0.5	0.046	0.090	0.999	0.999
		0.10	0.5	0.052	0.099	0.999	0.999
		0.25	0.5	0.049	0.077	0.998	0.998
		0.50	0.5	0.054	0.071	0.983	0.984
		0.75	0.5	0.049	0.062	0.949	0.951
		1.00	0.5	0.048	0.056	0.881	0.884

References

[1] Krishnamoorthy, K. and Guo, H., Assessing occupational exposure via the one-way random effects model with unbalanced data, *Journal of Statistical Planning and Inference*, **128**(2005), 219–229.

[2] Krishnamoorthy, K. and Mathew, T., Assessing occupational exposure via the one-way random effects model with balanced data, *Journal of Agricultural, Biological and Environment Statistics*, **7**(2002), 440–451.

[3] Li, X. and Li, G., Confidence intervals on sum of variance components with unbalanced designs, *Commun. Statist. – Theo.*, **34**(2005), 833–845.

[4] Lyles, R.V., Kupper, L.L. and Rapport, S.M., Assessing regulatory compliance of occupational exposures via the balanced one-way random effects ANOVA model, *Journal of Agricultural, Biological and Environment Statistics*, **2**(1997a), 64–86.

[5] Lyles, R.V., Kupper, L.L. and Rapport, S.M., A lognormal distribution-based exposure assessment method for unbalanced data, *Annals of Occupational Hygiene*, **41**(1997b), 63–76.

[6] Maxim, L.D., Allshouse, J.N. and Venturin, D.E., The random-effects model applied to refractory ceramic fiber data, *Regulatory Toxicology and Pharmacology*, **32**(2000), 190–199.

[7] Mu, W., Xiong, S. and Xu, X., Assessing occupational exposure via the two-way random effects model, *J. Sys. Sci. & Math. Scis.*, **27**(2007), 134–144.

[8] Rapport, S.M., Kromhout, H. and Venturine, D.E., Variation of exposure between worker in homogenous exposure groups, *American Industrial Hygiene Association Journal*, **54**(1993), 654–662.

[9] Rapport, S.M., Lyles, R.H. and Kupper, L.L., An exposure-assessment strategy accounting for within- and between-worker source of variability, *Annals of Occupational Hygiene*, **39**(1995), 469–495.

[10] Thomas, J.D. and Hultquist, R.A., Interval estimation for the unbalanced case of the one-way random effects model, *Ann. Statist.*, **6**(1978), 582–587.

[11] Tsui, K. and Weerahandi, S., Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters, *J. Amer. Statist. Assoc.*, **84**(1989), 602–607.

[12] Weerahandi, S., Generalized confidence intervals, *J. Amer. Statist. Assoc.*, **88**(1993), 899–905.

不平衡单因素随机模型在职业接触评价中的应用

李 新 民

(山东理工大学理学院应用数学研究所, 淄博, 255049)

在职业接触评价中, 单因素随机模型可用于评价工人接触均值超过职业接触限值的概率. 当数据不平衡时, 本文利用广义推断研究了关于此概率的假设检验, 并对此方法与已有方法进行了模拟对比研究. 模拟结果表明, 本文所给方法优于已有方法, 特别是在数据极不平衡时效果更优.

关键词: 广义p值, 势, 随机变量, 职业接触限值.

学科分类号: O212.6.