Tatra Mt. Math. Publ. **26** (2003), 175–182

TATRA MOUNTAINS
Mathematical Publications

INTERVAL ESTIMATORS FOR THE BETWEEN LABORATORY VARIANCE COMPONENT

Gejza Wimmer

ABSTRACT. The studied topic is motivated by the problem of interlaboratory comparisons. The paper presents a brief overview of the methods for deriving the interval estimators for the between laboratory variance under homoscedasticity. Under the heteroscedasticity we propose interval estimators based on the generalized p-values. Due to computational reasons a modified version of the generalized p-value estimator (MPV) is suggested.

1. Introduction

The problem of interlaboratory comparisons is of particular interest for many applications in industrial and scientific practice, see, e.g., Paule and Mandel (1982), Rukhin and Vangel (1998), Witkovský (2001a), and Witkovský and Wimmer (2002).

In particular, we consider measurements of the quantity of interest on virtually the same object which are made by k, $k \geq 2$, laboratories. The ith laboratory repeats its measurements n_i times, $n_i \geq 2$. The laboratories may exhibit the between laboratory variability, as well as different within laboratory variances (heteroscedasticity). We will assume that the measurements follow normal distribution.

In this paper we focus on the interval estimation of the between laboratory variance component in the one-way random effects model which characterizes the between laboratory variability, that is

$$Y_{ij} = \mu + a_i + \varepsilon_{ij}\,. \tag{1}$$

Here μ is the true (unknown) measured value of the quantity of interest, a_i are the random laboratory effects and $\varepsilon_{ij} \sim N(0, \sigma_i^2)$, $i = 1, \dots, k$, $j = 1, \dots, n_i$,

2000 Mathematics Subject Classification: 62F25.

Keywords: unbalanced heteroscedastic one-way random effects model, between laboratory variance, confidence interval, interlaboratory comparisons.

The research has been supported by the Grant VEGA 1/7295/20 from the Science Grant Agency of the Slovak Republic.

are mutually independent errors. The result of (i,j)th measurement is y_{ij}; the realization of the random variable Y_{ij}.

Further, we assume that $a_i \sim N(0,\sigma_a^2)$ are mutually independent and independent with all ε_{ij}. The variance components σ_i^2 and σ_a^2 are the nuisance parameters: the within laboratory and between laboratory variances.

2. Between laboratory variance interval estimation under homoscedasticity

Current state of the between laboratory variance (σ_a^2) interval estimation in the unbalanced one-way random effects model (1) with equal within laboratory variances (the homoscedastic case), i.e., $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$, is given in the paper by Hartung and Knapp (2000).

In this setup Thomas and Hultquist (1978) showed that

$$\frac{(k-1)\mathrm{MSA}}{\sigma_a^2 + \sigma^2/\tilde{n}} \overset{\text{approx.}}{\sim} \chi^2_{k-1}\,, \tag{2}$$

where

$$\mathrm{MSA} = \frac{1}{k-1}\sum_{i=1}^{k}(\bar{Y}_i - \bar{\bar{Y}})^2 \tag{3}$$

with $\bar{\bar{Y}} = \frac{1}{k}\sum\limits_{i=1}^{k}\bar{Y}_i$ and $\tilde{n} = k/\sum\limits_{i=1}^{k}\frac{1}{n_i}$. Adopting the ideas by Tukey (1951) and Williams (1962) they proposed the following approximate $(1-\alpha)$ confidence interval $\langle L_{TH}, U_{TH}\rangle$ for σ_a^2;

$$\begin{aligned} L_{TH} &= \frac{(k-1)}{\chi^2_{k-1}(1-\alpha/2)}\left(\mathrm{MSA} - \frac{F_{k-1,n-k}(1-\alpha/2)}{\tilde{n}}\,\mathrm{MSE}\right),\\ U_{TH} &= \frac{(k-1)}{\chi^2_{k-1}(\alpha/2)}\left(\mathrm{MSA} - \frac{F_{k-1,n-k}(\alpha/2)}{\tilde{n}}\,\mathrm{MSE}\right), \end{aligned} \tag{4}$$

with

$$\mathrm{MSE} = \frac{1}{n-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2 \tag{5}$$

and where $\chi^2_\nu(\gamma)$ denotes the γ-quantile of a χ^2-distribution with ν degrees of freedom, $F_{\nu_1,\nu_2}(\gamma)$ denotes the γ-quantile of a F-distribution with ν_1 and ν_2 degrees of freedom.

Burdick and Eickman (1986) investigated the approximate $(1-\alpha)$ confidence interval $\langle L_{BE}, U_{BE}\rangle$, where

$$\begin{aligned} L_{BE} &= \Big(\frac{\tilde{n}l}{1+\tilde{n}l}\Big)\frac{(k-1)}{\chi^2_{k-1}(1-\alpha/2)}\,\mathrm{MSA}\,,\\ U_{BE} &= \Big(\frac{\tilde{n}u}{1+\tilde{n}u}\Big)\frac{(k-1)}{\chi^2_{k-1}(\alpha/2)}\,\mathrm{MSA}\,, \end{aligned} \tag{6}$$

with

$$\begin{aligned} l &= \max\Big\{0, \frac{1}{F_{k-1,n-k}(1-\alpha/2)}\frac{\mathrm{MSA}}{\mathrm{MSE}} - \frac{1}{n_{\min}}\Big\}\,,\\ u &= \max\Big\{0, \frac{1}{F_{k-1,n-k}(\alpha/2)}\frac{\mathrm{MSA}}{\mathrm{MSE}} - \frac{1}{n_{\max}}\Big\}\,, \end{aligned} \tag{7}$$

$n_{\min} = \min\limits_{i=1,\dots,k}\{n_i\}$, $n_{\max} = \max\limits_{i=1,\dots,k}\{n_i\}$.

Let $\langle \eta_L, \eta_U\rangle$ be an exact $(1-\alpha)$ confidence interval for $\eta = \sigma_a^2/\sigma^2$ given in Wald (1940), see also Hartung and Knapp (2000). Hartung and Knapp proposed two approximate $(1-\alpha)$ confidence intervals $\langle L_{HK1}, U_{HK1}\rangle$ and $\langle L_{HK2}, U_{HK2}\rangle$ for σ_a^2 where

$$\begin{aligned} L_{HK1} &= \frac{n-k}{\chi^2_{n-k}(1-\alpha/2)}\,\mathrm{MSE}\,\eta_L\,,\\ U_{HK1} &= \frac{n-k}{\chi^2_{n-k}(\alpha/2)}\,\mathrm{MSE}\,\eta_U\,, \end{aligned} \tag{8}$$

and

$$\begin{aligned} L_{HK2} &= \mathrm{MSE}\,\eta_L\,,\\ U_{HK2} &= \mathrm{MSE}\,\eta_U\,. \end{aligned} \tag{9}$$

Hartung and Knapp (2000) performed a large simulation study where they investigated the properties of the above given approximate confidence intervals for σ_a^2.

3. Between laboratory variance interval estimator based on the generalized p-values under heteroscedasticity

The concept of generalized p-values has been introduced by Tsui and Weerahandi (1989). Consider an observable random vector X such that its distribution depends on the vector parameter (θ, ϑ), where θ is the parameter of interest and ϑ is a vector of the other nuisance parameters. Further, consider the problem of testing one-sided hypothesis

$$H_0\colon \theta \le \theta_0\,, \quad \text{against} \quad H_1\colon \theta > \theta_0\,, \tag{10}$$

where θ_0 is a prespecified value of θ. Let x be an observed value of X. Then an observed significance level for hypothesis testing is defined on the basis of a data-based generalized extreme region, a subset of the sample space, with x on its boundary. In order to define such an extreme region, a stochastic ordering of the sample space according to the possible values of θ is required. This could be accomplished by means of generalized test variables.

A random variable $T(X, x, \theta_0, \vartheta)$ is said to be a *generalized test variable* if it has the following properties:

1. $T_{\mathrm{obs}}(\theta_0) = T(x, x, \theta_0, \vartheta)$ does not depend on the unknown nuisance parameters.
2. If $\theta = \theta_0$ (θ is the true value of the parameter), the probability distribution of $T(X, x, \theta_0, \vartheta)$ is free of nuisance vector parameter ϑ.
3. For fixed x and ϑ, and for any given t, $\Pr\{T(X, x, \theta_0, \vartheta) \le t \,|\, \theta\}$ is a monotone function of θ.

   If $\Pr\{T(X, x, \theta_0, \vartheta) \le t \,|\, \theta\}$ is a nondecreasing function of θ, then the test variable $T(X, x, \theta_0, \vartheta)$ is said to be stochastically increasing in θ. If $\Pr\{T(X, x, \theta_0, \vartheta) \le t|\theta\}$ is a nonincreasing function of θ, then the test variable $T(X, x, \theta_0, \vartheta)$ is said to be stochastically decreasing in θ.

If $T(X, x, \theta_0, \vartheta)$ is a stochastically increasing test variable, then the subset of the sample space, $C_x(\theta_0) = \{x_*\colon T(x_*, x, \theta_0, \vartheta) \ge T_{\mathrm{obs}}(\theta_0)\}$, is said to be a generalized extreme region for testing H_0 against H_1 and $p = \sup\limits_{\theta \le \theta_0} \Pr\{X \in C_x(\theta_0)|\theta\}$ is said to be its generalized p-value for testing H_0. If $T(X, x, \theta_0, \vartheta)$ is stochastically increasing then $p = \Pr\{T(X, x, \theta_0, \vartheta) \ge T_{\mathrm{obs}}(\theta_0)|\theta = \theta_0\}$. Notice that if $T(X, x, \theta_0, \vartheta)$ is stochastically decreasing then the p-value is $p = \Pr\{T(X, x, \theta_0, \vartheta) \le T_{\mathrm{obs}}(\theta_0)\,|\theta = \theta_0\}$.

If the null hypothesis is right-sided, then the generalized p-value for testing H_0 is $p = \Pr\{T(X, x, \theta_0, \vartheta) \le T_{\mathrm{obs}}(\theta_0)|\theta = \theta_0\}$, provided that $T(X, x, \theta_0, \vartheta)$ is stochastically increasing, and $p = \Pr\{T(X, x, \theta_0, \vartheta) \ge T_{\mathrm{obs}}(\theta_0)|\theta = \theta_0\}$, if $T(X, x, \theta_0, \vartheta)$ is stochastically decreasing. Based on the above discussion the concept of generalized p-values can be easily generalized for testing the two-sided null hypothesis $H_0\colon \theta = \theta_0$ against $H_1\colon \theta \ne \theta_0$, see also Witkovský (2001c).

It is easy to see that

$$\begin{aligned} SS &= Y'R'(V^{-1} - V^{-1}\mathit{1}(\mathit{1}'V^{-1}\mathit{1})^{-1}\mathit{1}'V^{-1})RY \\ &= \sum_{i=1}^{k} \frac{1}{\sigma_a^2 + \sigma_i^2/n_i} \left(\bar{Y}_i - \frac{\sum_{j=1}^{k} \frac{\bar{Y}_j}{\sigma_a^2+\sigma_j^2/n_j}}{\sum_{j=1}^{k} \frac{1}{\sigma_a^2+\sigma_j^2/n_j}} \right)^2 \sim \chi^2_{k-1}\,, \end{aligned} \tag{11}$$

where $R = (Z'Z)^{-1}Z'$ with $Z = \mathrm{Diag}\{\mathit{1}_i\}$ ($\mathit{1}_i$ is an n_i-vector of ones), $V = \sigma_a^2 I_k + \mathrm{Diag}\{\sigma_i^2/n_i\}$, and

$$\frac{n_i - 1}{\sigma_i^2} S_i^2 \sim \chi^2_{n_i-1}\,, \qquad i = 1, \ldots, k\,. \tag{12}$$

Moreover, these random variables are mutually independent. To construct the generalized confidence interval estimate for σ_a^2 we propose the following potential test variable

$$T = \frac{Y'R'(V^{-1} - V^{-1}\mathit{1}(\mathit{1}'V^{-1}\mathit{1})^{-1}\mathit{1}'V^{-1})RY}{y'R'(W^{-1} - W^{-1}\mathit{1}(\mathit{1}'W^{-1}\mathit{1})^{-1}\mathit{1}'W^{-1})Ry}\,, \tag{13}$$

where W denotes a matrix with random elements,

$$W = \sigma_a^2 I_k + \mathrm{Diag}\left\{ \frac{\sigma_i^2}{n_i - 1} \frac{n_i - 1}{n_i} \frac{s_i^2}{S_i^2} \right\} \tag{14}$$

and W is distributed as

$$\sigma_a^2 I_k + \mathrm{Diag}\left\{ \frac{n_i - 1}{n_i} \frac{s_i^2}{\chi^2_{n_i-1}} \right\}. \tag{15}$$

Note that the potential test variable fulfills the required properties 1–3, and the observed value $T_{\mathrm{obs}} = 1$. Given σ_a^2, the generalized p-value is

$$p = P(T > T_{\mathrm{obs}} | \sigma_a^2) = P\big(\chi^2_{k-1} > q(\sigma_a^2, \chi^2_{n_1-1}, \ldots, \chi^2_{n_k-1})\big)\,, \tag{16}$$

where $q(\sigma_a^2, \chi^2_{n_1-1}, \ldots, \chi^2_{n_k-1}) = y'R'(W^{-1} - W^{-1}\mathit{1}(\mathit{1}'W^{-1}\mathit{1})^{-1}\mathit{1}'W^{-1})Ry$ denotes random variable which depends on σ_a^2 and on the independent random variables with $\chi^2_{n_i-1}$ distribution, $i = 1, \ldots, k$.

The generalized $(1 - \alpha)$ confidence interval is then given as $\langle L_{PV}, U_{PV} \rangle$, where L_{PV} and U_{PV} are solutions of the following equations

$$\begin{aligned} \frac{\alpha}{2} &= P\big(\chi^2_{k-1} > q(L_{PV}, \chi^2_{n_1-1}, \ldots, \chi^2_{n_k-1})\big)\,, \\ 1 - \frac{\alpha}{2} &= P\big(\chi^2_{k-1} > q(U_{PV}, \chi^2_{n_1-1}, \ldots, \chi^2_{n_k-1})\big)\,. \end{aligned} \tag{17}$$

The exact evaluation of the required generalized p-values in general requires k-dimensional numerical integration. However, for practical applications it is possible to evaluate the p-values approximately by Monte Carlo simulations combined with the bisection method. For that, one needs to generate the realizations of independent random variables $\chi^2_{k-1}, \chi^2_{n_1-1}, \dots, \chi^2_{n_k-1}$ and count the relative frequencies of the cases in which $\chi^2_{k-1} > q(\sigma^2_a, \chi^2_{n_1-1}, \dots, \chi^2_{n_k-1})$ for any chosen value σ^2_a.

Here we suggest to consider an approximate $(1-\alpha)$ confidence interval (based on the modified generalized p-values). The lower and upper bounds of such interval can be calculated by one-dimensional numerical integration as the quantiles of the distribution of a linear combination of independent inverted chi-square random variables. For more details see W i t k o v s k ý (2001b), W i t k o v s k ý (2002).

W i m m e r and W i t k o v s k ý (2002) proved that

$$W = \frac{(k-1)\mathrm{MSA}}{\sigma^2_a + \frac{1}{k}\sum\limits_{j=1}^{k} \frac{\sigma^2_j}{n_j}} \overset{\text{approx.}}{\sim} \chi^2_{k-1}\,. \tag{18}$$

Based on that, the approximate $(1-\alpha)$ modified generalized p-value confidence interval for σ^2_a is defined as $\langle L_{MPV}, U_{MPV}\rangle$, where L_{MPV} and U_{MPV} are the solutions of the equations

$$\begin{aligned} \frac{\alpha}{2} &= P\left(\frac{\sum\limits_{i=1}^{k}(\bar{y}_i - \bar{\bar{y}})^2}{\chi^2_{k-1}} - \sum_{j=1}^{k}\left(\frac{(n_j-1)s^2_j}{kn_j}\right)\frac{1}{\chi^2_{n_j-1}} \le L_{MPV}\right), \\ 1-\frac{\alpha}{2} &= P\left(\frac{\sum\limits_{i=1}^{k}(\bar{y}_i - \bar{\bar{y}})^2}{\chi^2_{k-1}} - \sum_{j=1}^{k}\left(\frac{(n_j-1)s^2_j}{kn_j}\right)\frac{1}{\chi^2_{n_j-1}} \le U_{MPV}\right), \end{aligned} \tag{19}$$

where χ^2_{k-1} and $\chi^2_{n_j-1}$ are mutually independent chi-square distributed random variables with with $k-1$ and n_j-1 degrees of freedom, $j = 1, \dots, k$.

An intensive simulation study was made in order to compare the confidence coefficients and the expected lengths of several different confidence intervals for σ^2_a. The complete results of simulations are available on the request from the author and are also presented as Technical Report available at the web page `http://www.mat.savba.sk` (see the Preprint Series Section).

4. Example

Eberhardt *et al.* (1989) reported the data on Selenium in non-fat milk powder. The measurements in Table 1 are based on four independent measurement methods.

TABLE 1. Selenium in non-fat milk powder data. Source: Eberhardt *et al.* (1989).

Method	n_i	$\bar{y}_i$	s_i^2
A	8	105.00	85.711
B	12	109.75	20.748
C	14	109.50	2.729
D	8	113.25	33.640

For illustration purposes we have calculated the 95 % interval estimates for the between laboratory variance σ_a^2 by the methods mentioned in this paper, see Table 2.

TABLE 2. The 95 % interval estimates for the between laboratory variance σ_a^2.

Estimator	Equation	Lower bound	Upper bound
TH	(4)	0.3990	156.1074
BE	(6)	0.0000	156.1228
HK1	(8)	0.0000	260.7897
HK2	(9)	0.0000	157.0124
PV	(17)	0.0000	148.4149
MPV	(19)	0.0000	153.4232

REFERENCES

[1] BURDICK, R. K.—EICKMAN, J.: *Confidence intervals on the among laboratory variance component in the unbalanced one-fold nested design*, J. Statist. Comput. Simulation **26** (1986), 205–219.

[2] EBERHARDT, K. R.—REEVE, C. P.—SPIEGELMAN, C. H.: *A minimax approach to combining means, with practical examples*, Chemometrics Intell. Lab. Systems **5** (1989), 129–148.

[3] HARTUNG, J.—KNAPP, G.: *Confidence intervals for the between group variance in the unbalanced one-way random effects model of analysis of variance*, J. Statist. Comput. Simulation **65** (2000), 311–323.

[4] PAULE, R. C.—MANDEL, J.: *Consensus values and weighting factors*, J. Res. Nat. Bur. Stand. **87** (1982), 377–385.

[5] RUKHIN, A. L.—VANGEL, M. G.: *Estimation of a common mean and weighted means statistics*, J. Amer. Statist. Assoc. **93** (1998), 303–308.

[6] THOMAS, J. D.—HULTQUIST, R. A.: *Interval estimation for the unbalanced case of the one-way random effects model*, Ann. Statist. **6** (1978), 582–587.

[7] TSUI, K. W.—WEERAHANDI, S.: *Generalized p values in significance testing of hypotheses in the presence of nuisance parameters*, J. Amer. Statist. Assoc. **84** (1989), 602–607.

[8] TUKEY, J. W.: *Components in regression*, Biometrics **7** (1951), 33–69.

[9] WALD, A.: *A note on the analysis of variance with unequal class frequencies*, Ann. Math. Stat. **11** (1940), 96–100.

[10] WILLIAMS, J. S.: *A confidence interval for variance components*, Biometrika **49** (1962), 278–281.

[11] WIMMER, G.—WITKOVSKÝ, V.: *Between group variance component interval estimation for the unbalanced heteroscedastic one-way random effects model*, J. Statist. Comput. Simulation **73** (2003), 333–346.

[12] WITKOVSKÝ, V.: *On the exact computation of the density and of the quantiles of linear combinations of t and F random variables*, J. Statist. Plann. Inference **94** (2001a), 1–13.

[13] WITKOVSKÝ, V.: *Computing the distribution of a linear combination of inverted gamma variables*, Kybernetika **37** (2001b), 79–90.

[14] WITKOVSKÝ, V.: *Exact tests of variance components using generalized p-values*, Folia Fac. Sci. Nat. Univ. Masarykianae Brunensis, Mathematica **9** (2001c), 119–125.

[15] WITKOVSKÝ, V.: *Exact distribution of positive linear combinations of inverted chi-square random variables with odd degrees of freedom*, Statist. Probab. Lett. **56** (2002), 45–50.

[16] WITKOVSKÝ, V.—WIMMER, G.: *Consensus mean and interval estimators for the common mean*, Tatra Mt. Math. Publ. **26** (2003), 183–194.

Received May 3, 2002

Faculty of Natural Sciences
Matej Bel University
Tajovského 40
SK–974 00 Banská Bystrica
SLOVAKIA

Mathematical Institute
Slovak Academy of Sciences
Šefánikova 49
SK–814 73 Bratislava
SLOVAKIA

E-mail: wimmer@mat.savba.sk