Comparison of some Exact and Approximate Interval Estimators for Common Mean[1]

Viktor Witkovský

Institute of Measurement Science, Slovak Academy of Sciences
Dúbravská cesta 9, 841 04 Bratislava, Slovak Republic
E-mail: witkovsky@savba.sk

Abstract. *In this paper we compare several exact and approximate interval estimators, suggested in recent statistical literature, for the common mean in the one-way fixed effects model. The the common mean problem is also known in metrology where the experiments based on interlaboratory comparisons are provided in order to identify the comparison reference value (CRV).*

Keywords: *Interlaboratory studies; Comparison Reference Value; Common mean; Confidence Interval.*

1. Introduction

The Mutual Recognition Arrangement (MRA), see [1], prompted research of suitable statistical procedures to summarize the results of interlaboratory comparisons (IC). In this paper we consider particular type of IC, such that all participating measurement institutes (typically the National Measurement Institutes, NMIs) measure (unbiasedly) the same standard, and no other uncertainty could influence the measurement process except the measurement errors of the participating NMIs. From statistical point of view the problem of deriving the comparison reference value (CRV) is equivalent to the problem known as the common mean problem. In this paper we will present and compare several exact and approximate interval estimators, suggested in recent statistical literature, for the common mean in the one-way fixed effects model.

We will consider the following model:

$$y_{ij} = \mu + \varepsilon_{ij}, \tag{1}$$

with mutually independent errors, distributed according to normal (gaussian) distribution, i.e. $\varepsilon_{ij} \sim N(0, \sigma_i^2)$, $i = 1, \dots, k$, and $j = 1, \dots, n_i$. The variance components σ_i^2 are the nuisance parameters, which could be, in general, unequal. The outcome of the IC experiment is typically given by the estimates of NMIs means and error variances. We will use the following notation: $\bar{y}_i = (1/n_i)\sum_{j=1}^{n_i} y_{ij}$, $s_i^2 = (1/(n_i - 1))\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2$.

Under the assumption that IC measurements follow the model (1) and if the variance components σ_i^2 are known, the optimal estimator for the unknown common mean μ would be the generalized least squares estimator $\hat{\mu}_{GLS} = (\sum_{i=1}^{k}\omega_i\bar{y}_i)/(\sum_{i=1}^{k}\omega_i)$, where $\omega_i = n_i/\sigma_i^2$. The exact distribution of the

[1]The research was supported by grant from Scientific Grant Agency of the Slovak Republic VEGA 1/0264/03.

estimator is known: $\hat{\mu}_{GLS} \sim N(\mu, 1/\omega_\Sigma)$, where $\omega_\Sigma = \sum_{i=1}^{k} \omega_i$. From that the corresponding exact $(1-\alpha) \times 100\%$ confidence interval is given by

$$CI_1 : \left[\hat{\mu}_{GLS} - u_{1-\alpha/2}\sqrt{1/\omega_\Sigma},\ \hat{\mu}_{GLS} + u_{1-\alpha/2}\sqrt{1/\omega_\Sigma}\right], \tag{2}$$

where $u_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the standard normal distribution $N(0,1)$. This confidence interval is based on full information about the model and the nuisance parameters σ_i^2, $i = 1, \ldots, k$, which is in typical applications unknown. However, the interval estimator(2) will serve as a gold standard (benchmark) in further analysis of statistical properties of the other suggested interval estimators for the common mean.

If the variance components σ_i^2 are unknown (only the estimates s_i^2 are available) the situation is more complicated, and the exact and optimum (shortest) interval estimator for the common mean μ is not known. The problem considered here was studied by many authors in the statistical literature, see e.g. [2, 4, 5, 8, 9]. This paper is also related to the work presented by Savin in [6]. In the next Section we present some approximate and exact confidence intervals for the common mean which have been proved to have good statistical properties under different situations.

2. Exact and approximate confidence intervals for the common mean

Fairweather, see [2], suggested the exact confidence interval for the common mean μ based on the distribution of a linear combination of independent Student's t random variables, say $W = \sum_{i=1}^{k} u_i t_i$, where u_i, $i = 1, \ldots, k$, denote the nonstochastic coefficients which represent the relative importance of the participating NMIs, and t_i, $i = 1, \ldots, k$, denote the independent Student's t random variables with $\nu_i = n_i - 1$ degrees of freedom. If we denote by $q_{1-\alpha/2}$ the $(1-\alpha/2)$-quantile of the distribution of W, then the exact $(1-\alpha) \times 100\%$ confidence interval for μ is given by

$$\left[\frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i} - \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i},\ \frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i} + \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}u_i}\right]. \tag{3}$$

The quantiles of W could be calculated exactly by the algorithm `tdist` suggested by Witkovský in [7].

Here we consider two versions of confidence interval (3) based on different choice of weights u_i, $i = 1, \ldots, k$: If we have no reason to prefer results of any particular participating NMI, we suggest to use the confidence interval (3) with $u_i = 1$ for all $i = 1, \ldots, k$, i.e.

$$CI_2 : \left[\frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}} - \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}},\ \frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}} + \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}}\right], \tag{4}$$

where $q_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the distribution of $W = \sum_{i=1}^{k} t_i$.

On the other hand, if we have some prior information on the NMIs error variances (prior to the current IC experiment), say, we know from preliminary experiments or from the analysis of the measurement devices that the error variances are $\sigma_{(0)i}^2$, we suggest to use the confidence interval (3) with $u_i = \sqrt{n_i/\sigma_{(0)i}^2}$ for all $i = 1, \ldots, k$, i.e.

$$CI_3 : \left[\frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}} - \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}},\ \frac{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}\bar{y}_i}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}} + \frac{q_{1-\alpha/2}}{\sum_{i=1}^{k}\sqrt{\frac{n_i}{s_i^2}}\sqrt{\frac{n_i}{\sigma_{(0)i}^2}}}\right], \tag{5}$$

where $q_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the distribution of $W = \sum_{i=1}^{k} \sqrt{n_i/\sigma^2_{(0)i}} t_i$.

Hartung and Makambi in [4] suggested the following approximate confidence intervals for μ, centered at the Graybill-Deal estimator $\hat{\mu}_{GD}$ of μ, see [3]:

$$CI_4 : \left[\hat{\mu}_{GD} - t_{\hat{\nu};1-\alpha/2}\sqrt{1/w_\Sigma},\ \hat{\mu}_{GD} + t_{\hat{\nu};1-\alpha/2}\sqrt{1/w_\Sigma}\right], \tag{6}$$

where $\hat{\mu}_{GD} = (\sum_{i=1}^{k} w_i\bar{y}_i)/(\sum_{i=1}^{k} w_i)$ with $w_i = n_i/s_i^2$, $w_\Sigma = \sum_{i=1}^{k} w_i$, and $t_{\hat{\nu};1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the Students t-distribution with degrees of freedom estimated by $\hat{\nu} = (2f)/(f-1)$, where

$$f = 1 + \frac{2}{w^2_{\Sigma^*}} \sum_{i=1}^{k} \frac{w_i}{n_i - 1}(2w_\Sigma - w_i), \quad \text{and} \quad w_{\Sigma^*} = \sum_{i=1}^{k} \frac{(n_i-3)}{(n_i-1)} \frac{n_i}{s_i^2}. \tag{7}$$

The other approximate confidence interval for μ centered at $\hat{\mu}_{GD}$, suggested by Hartung and Makambi in [4], is the confidence interval given by

$$CI_5 : \left[\hat{\mu}_{GD} - t_{\hat{\nu}_*;1-\alpha/2}\sqrt{\hat{\lambda}/w_\Sigma},\ \hat{\mu}_{GD} + t_{\hat{\nu}_*;1-\alpha/2}\sqrt{\hat{\lambda}/w_\Sigma}\right], \tag{8}$$

with degrees of freedom estimated by $\hat{\nu}_* = 4 + (6f^2)/|V - 2f^2|$, where f is defined in (7), and

$$V = \frac{2}{w^2_\Sigma}\left\{\sum_{i=1}^{k} w_i^2\left(1 + \frac{14}{n_i-1}\right) - \frac{8}{w_\Sigma}\sum_{i=1}^{k}\frac{w_i^3}{n_i-1}\right\}, \quad \hat{\lambda} = \frac{\hat{\nu}_*}{(\hat{\nu}_*-2)f}. \tag{9}$$

Krishnamoorthy and Lu, see [5], suggested an approximate confidence interval based on the generalized pivot, which is a stochastic linear combination of Student's t random variables:

$$T = \frac{\sum_{i=1}^{k} \frac{n_iQ_i}{(n_i-1)s_i^2}(\bar{y}_i - \sqrt{s_i^2/n_i}t_i)}{\sum_{i=1}^{k} \frac{n_iQ_i}{(n_i-1)s_i^2}}, \tag{10}$$

where $Q_i \sim \chi^2_{\nu_i}$ are stochastically independent random variables distributed as chi-squared random variables with $\nu_i = n_i - 1$ degrees of freedom, which are also independent with $t_i \sim t_{\nu_i}$. The $(1-\alpha)\times 100\%$ confidence interval for μ is given by

$$CI_6 : \left[T_{\alpha/2},\ T_{1-\alpha/2}\right], \tag{11}$$

where $T_{\alpha/2}$ and $T_{1-\alpha/2}$ are the quantiles of the distribution of the random variable T given in (10). The quantiles of T could be estimated by an auxiliary simulation experiment.

3. Simulation study

We have examined the empirical coverage probabilities and the relative average lengths (relative to the average lengths of CI_1) of the interval estimators $CI_1, \ldots, CI_6$, for nominal level $\alpha = 0.05$. Assuming that model (1) is true, we have used the following values of the unknown parameters in the simulation study: $\mu = 0$, $k = 9$ and the following 10 sample designs used by Hartung and Makambi in [4]: 1) $n_i \in \{\{10,10,10\},\{10,10,10\},\{10,10,10\}\}$ and $\sigma_i^2 \in \{\{4,4,4\},\{4,4,4\},\{4,4,4\}\}$, 2) $n_i \in \{\{10,10,10\},\ldots\}$ and $\sigma_i^2 \in \{\{1,3,5\},\ldots\}$, 3) $n_i \in \{\{20,20,20\},\ldots\}$ and $\sigma_i^2 \in \{\{4,4,4\},\ldots\}$, 4) $n_i \in \{\{20,20,20\},\ldots\}$ and $\sigma_i^2 \in \{\{1,3,5\},\ldots\}$, 5) $n_i \in \{\{5,10,15\},\ldots\}$ and $\sigma_i^2 \in \{\{4,4,4\},\ldots\}$, 6) $n_i \in \{\{5,10,15\},\ldots\}$ and $\sigma_i^2 \in \{\{1,3,5\},\ldots\}$, 7) $n_i \in \{\{5,10,15\},\ldots\}$ and $\sigma_i^2 \in \{\{5,3,1\},\ldots\}$, 8)

Design	CI_1	CI_2	CI_3	CI_4	CI_5	CI_6
1	0.9490	0.9501 (1.05)	0.9501 (1.05)	0.9536 (1.18)	0.9542 (1.18)	0.9618 (1.21)
2	0.9497	0.9478 (1.11)	0.9474 (1.05)	0.9493 (1.18)	0.9500 (1.18)	0.9572 (1.20)
3	0.9494	0.9511 (1.02)	0.9511 (1.02)	0.9537 (1.08)	0.9541 (1.08)	0.9586 (1.10)
4	0.9472	0.9474 (1.08)	0.9473 (1.02)	0.9494 (1.08)	0.9496 (1.08)	0.9541 (1.10)
5	0.9516	0.9522 (1.13)	0.9512 (1.07)	0.9430 (1.20)	0.9462 (1.22)	0.9612 (1.27)
6	0.9478	0.9488 (1.09)	0.9492 (1.11)	0.9448 (1.29)	0.9539 (1.35)	0.9618 (1.32)
7	0.9478	0.9472 (1.27)	0.9453 (1.04)	0.9371 (1.15)	0.9375 (1.15)	0.9561 (1.21)
8	0.9497	0.9515 (1.06)	0.9512 (1.02)	0.9535 (1.08)	0.9540 (1.09)	0.9593 (1.11)
9	0.9472	0.9470 (1.03)	0.9448 (1.03)	0.9495 (1.12)	0.9493 (1.12)	0.9565 (1.14)
10	0.9497	0.9506 (1.18)	0.9481 (1.02)	0.9522 (1.06)	0.9537 (1.07)	0.9574 (1.08)

Table 1: Empirical coverage probabilities of nominal 95% confidence intervals $CI_1, \ldots, CI_6$ and their relative (with respect to CI_1) average lengths.

$n_i \in \{\{10, 20, 30\}, \ldots\}$ and $\sigma_i^2 \in \{\{4, 4, 4\}, \ldots\}$, 9) $n_i \in \{\{10, 20, 30\}, \ldots\}$ and $\sigma_i^2 \in \{\{1, 3, 5\}, \ldots\}$, and 10) $n_i \in \{\{10, 20, 30\}, \ldots\}$ and $\sigma_i^2 \in \{\{5, 3, 1\}, \ldots\}$. For each particular design we have generated 10000 realizations of the particular interval estimator. Each realization of CI_6 was based on another 10000 realizations of the pivotal quantity T given by (10). The simulations show that the confidence intervals $CI_2, \ldots, CI_6$ have good coverage properties for considered experimental designs. The empirical coverage probabilities of the approximate confidence intervals CI_4, CI_5, and CI_6 fluctuates around the nominal level in the range $\pm 1.5\%$. The average lengths exceed the expected length of the CI_1 in the range $1.02 - 1.35$.

References

[1] CIPM (1999). *Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates Issued by National Institutes.* Sèvres: Bureau International des Poids et Mesures.

[2] W.R. Fairweather (1972). A method of obtaining an exact confidence interval for the common mean of several normal populations. *Applied Statistics*, 21, 229–233.

[3] F.A. Graybill and R.D. Deal (1959). Combining unbiased estimators. *Biometrics*, 3, 1–21.

[4] J. Hartung and K.H. Makambi (2000). Alternative test procedures and confidence intervals on the common mean in the fixed effects model for meta-analysis. *Technical Report TR46/00*, University of Dortmund. `http://www.sfb475.uni-dortmund.de/berichte/tr46-00.ps`.

[5] K. Krishnamoorthy and Y. Lu (2003). Inferences on the common mean of several normal populations based on the generalized variable method. *Biometrics*, 59, 237–247.

[6] A. Savin (2005). Confidence intervals for common mean in one-way classification model with fixed effects. *MEASUREMENT 2005*. Submitted.

[7] V. Witkovský (2004). Matlab algorithm TDIST: The distribution of a linear combination of Student's t random variables. *In Jaromir Antoch (Ed.): COMPSTAT. Proceedings in Computational Statistics. 16th Symposium Held in Prague, Czech Republic, 2004*, Physica-Verlag, 1995–2002. See also `http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=4199&objectType=FILE`

[8] V. Witkovský and G. Wimmer (2003). Consensus mean and interval estimators for the common mean. *Tatra Mountains Mathematical Publications*, 26, 183–194.

[9] L.H. Yu, Y. Sun and B.K. Sinha (1999). On exact confidence intervals for the common mean of several normal populations. *Journal of Statistical Planning and Inference* 81, 263–277.