Comparison of a random effects model and a Birge ratio method for the adjustment of inconsistent data

Alfred Link, Olha Bodnar and Clemens Elster
Physikalisch-Technische Bundesanstalt, Braunschweig and Berlin, Germany

Abstract

We compare a random effects model and a variant of the Birge ratio method [1], [2] for adjusting inconsistent data. This task is relevant in metrology, for instance, in the agreement on values for fundamental constants. Although random effects models represent a standard tool in statistics (e.g. [3]) and their application in metrology has been advocated (e.g. [4], [5] [6]) they do not seem to be very popular among metrologists so far, in contrast to the Birge ratio method.

For the random effects model we assume that the given data is distributed as $\mathbf{x}|\mathbf{b} \sim \mathbf{N}(\mu\mathbf{1}+\mathbf{b}, \mathbf{V})$ with $\mathbf{b} \sim \mathbf{N}(\mathbf{0}, \sigma_\mathbf{b}^\mathbf{2}\mathbf{I})$. The symmetric positive definite uncertainty matrix $\mathbf{V}$ is assumed to be known. The goal is to infer μ given the observations $x_1, \ldots, x_n$, and we employ a Bayesian approach. In contrast to typical applications of random effects models no replicated within group measurements are available, and also the given matrix $\mathbf{V}$ might be seen as specific to metrology. We derive the Berger & Bernardo reference prior [7] in the marginal model $\mathbf{x} \sim \mathbf{N}(\mu\mathbf{1}, \mathbf{V} + \sigma_\mathbf{b}^\mathbf{2}\mathbf{I})$ for the group ordering $\{\mu, \sigma_b\}$. The Birge ratio method assumes the model $\mathbf{x} \sim \mathbf{N}(\mu\mathbf{1}, \mathbf{c}^\mathbf{2}\mathbf{V})$ with unknowns $\{\mu, c\}$. Also for this model we derive the Berger & Bernardo reference prior and give the resulting posterior for μ in analytical form. We note that our variant of the Birge ratio method produces more conservative uncertainties than the conventional Birge ratio method, in particular for small n. We argue that this is an improvement, cf. also [6].

We then compare the two methods and explore their robustness by analyzing simulated data. The simulations are based on data drawn from each of the two statistical models so that each method is once applied when its assumptions are met, and once when this is not the case. In addition we also study the robustness of the methods when the normality assumptions are violated. Finally, we discuss results obtained for the example of the adjustment of the Planck constant from current measurements.

References

[1] R T Birge 1932. The calculation of errors by the method of least squares *Phys. Rev*, 40, 207-227.

[2] P J Mohr and B N Taylor 2000. CODATA recommended values of the fundamental physical constants: 1998 *Reviews of Modern Physics*, 72, 351-495.

[3] J C Pinheiro and D M Bates 2000. *Mixed-Effects Models in S and S-PLUS* Springer, New York.

[4] R N Kacker 2004. Combining information from interlaboratory evaluations using a random effects model *Metrologia*, 41 132136.

[5] B Toman and A Possolo 2009. Laboratory effects models for interlaboratory comparisons *Accreditation and Quality Assurance*, 14, 553-563.

[6] B Toman, J Fischer and C Elster 2012. Alternative analyses of measurements of the Planck constant *Metrologia*, 49, 567-571.

[7] J Berger and J M Bernardo 1992 Ordered group reference priors with application to the multinomial problem *Biometrika*, 79, 25-37.