Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Rejoinder on the normalized maximum likelihood and Bayesian decision theory: Reply to Grünwald and Navarro (2009)[☆]

George Karabatsos[a,*], Stephen G. Walker[b]

[a] University of Illinois-Chicago, 1040 W. Harrison Street (MC 147), Chicago, IL 60607, United States

[b] University of Kent, Institute of Mathematics, Statistics and Actuarial Science, Canterbury, CT2 7NZ, United Kingdom

ARTICLE INFO

Article history:
Received 7 December 2008
Available online 13 January 2009

Keywords:
Bayesian nonparametrics
Normalized maximum likelihood
Model selection

ABSTRACT

We provide a short reply to [Grünwald, P., & Navarro, D. (2009). NML, Bayes and true distributions: A comment on Karabatsos and Walker. Journal of Mathematical Psychology, in press (doi:10.1016/j.jmp.2008.11.005)] comment on the article by [Karabatsos, G., & Walker, S. (2006). On the normalized maximum likelihood and Bayesian decision theory. Journal of Mathematical Psychology, 50, 517–520].

Grünwald and Navarro (2009) appear to have difficulty relating to the paper (Karabatsos & Walker, 2006) in which we discussed aspects of the Normalized Maximum Likelihood (NML) criterion for model selection. In particular, using straightforward algebra, we demonstrated in Karabatsos and Walker (2006) that the NML criterion for model selection can be characterized as an approach to Bayesian model selection under a noninformative (Dirichlet Process) prior and under a specific logarithmic utility function which incorporates a particular penalty for model complexity.

So let us write it as simply as we can. A well known loss function when using the density function $f(x;\theta)$ and where P is the true distribution function is given by

$$l(\theta, P) = -\int \log f(x;\theta)\,\mathrm{d}P(x).$$

Using standard maximum of expected utility theory, and using a Dirichlet process prior, given observations $(x_1,\ldots,x_n)$, we have a posterior expected utility of

$$U(\theta) = \int \log f(x;\theta)\left\{\frac{c\mathrm{d}P_0(x) + n\mathrm{d}G_n(x)}{c+n}\right\},$$

where P_0 is the prior guess at P, $c > 0$ is a scale parameter and G_n the empirical c.d.f. (i.e., the distribution function for the n observations). So,

$$U(\theta) = w_n \int \log f(x;\theta)\,\mathrm{d}P_0(x) + (1-w_n)\,n^{-1}\sum_{i=1}^{n} \log f(x_i,\theta),$$

where $w_n = c/(c+n)$. Letting c be very small, so small that it can be ignored or put to 0, then the θ which maximizes $U(\theta)$ is the maximum likelihood estimator. We are not sure what we would do differently here if we had the luxury to philosophize about if a true P existed or not.

Now we decide to add some penalty to this loss function for the dimension of θ which depends on the sample size — call it $v(d(\theta), n)$. So we now want to maximize

$$U(\theta) = \sum_{i=1}^{n} \log f(x_i,\theta) - v(\mathrm{d}(\theta), n).$$

Our simple point, now, is that for a particular choice of $v(d,n)$ we recover NML.

The Editor invited us to give a short reply to the article by Grünwald and Navarro (2009), and we thank him for this. However, we were not sure what could be said other than what we wrote in Karabatsos and Walker (2006) is quite straightforward, and that not only did we find Grünwald and Navarro's (2009) article to be a bit misleading, but also that they have, by and large, misunderstood and misrepresented our article. Furthermore, we do not see how the arguments raised by Grünwald and Navarro (2009) make any of Karabatsos and Walker's (2006) conclusions false. So we respectfully decline to comment any further on their paper.

References

Grünwald, P., & Navarro, D. (2009). NML, Bayes and true distributions: A comment on Karabatsos and Walker. *Journal of Mathematical Psychology*, *53*(1), 43–51.

Karabatsos, G., & Walker, S. (2006). On the normalized maximum likelihood and Bayesian decision theory. *Journal of Mathematical Psychology*, *50*, 517–520.

☆ This research is supported by National Science Foundation grant SES-0242030 in the program of Methodology, Measurement, and Statistics. The views expressed in this article reflects those of the authors.

* Corresponding author.
E-mail addresses: georgek@uic.edu (G. Karabatsos), S.G.Walker@kent.ac.uk (S.G. Walker).

0022-2496/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2008.12.003