Discussion of Bayesian Synthetic Likelihood – Asymptotics and Misspecification

Stefano Cabras, Universidad Carlos III de Madrid (Spain)

Santa Cruz (CA) - September 8^{th}, 2022

Prologue

- I make the Discussion based on these three recommended papers:
  1. Frazier et al. (2022);
  2. Frazier and Drovandi (2021) and
  3. Frazier, Drovandi, and Nott (2021)
- Authors made contributions on the *aymptotic* behaviour of posterior in the context of *unavailable likelihood* expressions:
  - This is totally new.
  - We need these studies.
  - Great job!

Summary

1. What I understand about BSL.
2. Similarities of BSLs with other SL.
3. Some words on misspecification: parametrization/parameter interpretation.

Setup and definition of BSL

- Consider the model $\pi(y \mid \theta)$ with unknown marginal likelihood for $\theta \in \mathbb{R}^d$, $d \geq 1$;
- $s = s(y) \in \mathbb{R}$ is a summary statistic, generated from $\pi(s(y) \mid \theta)$, where:
  - the observed value from a sample y_n is s_{obs};
  - the BSL is

$$L_{BSL}(\theta) := \phi\left(\frac{b(\theta) - s_{obs}}{\sigma_R(\theta)}\right),$$

- where:
  - $b(\theta) = E(S \mid \theta)$ is approximated by MC on m simulations of s at each θ from $\pi(s(y)|\theta)$;
  - $\sigma_R(\theta)$ is the Variance / Covariance matrix (for $d > 1$);
    - different estimations are considered to deal when no θ makes $b(\theta) - s_{obs} = 0$ (model Missspecification).

Other versions of the BSL? : Quasi Likelihood (QL)

- Authors mentioned other ways to replace the usual ABC likelihood:
  - Empirical likelihood, as in Chaudhuri et al. (2020)
  - in general: adjustments of usual ABC output

- I would like to mention among these also the Quasi-Likelihood (QL) $L_Q(\boldsymbol{\theta})$:
  - defined upon the theory of estimating equations (see McCullagh (1991));
- Let $\Psi = \Psi(s;\theta)$ be an unbiased estimating function such that $E_{\pi(s(y)|\theta)}(\Psi \mid \theta) = 0$, then

$$L_Q(\theta) = \exp\left\{\int_{c_0}^{\theta} A(t)\Psi(s;t)\,dt\right\}$$

- where
  - c_0 is an arbitrary constant;
  - $A(\theta) = B(\theta)/\Omega(\theta)$;
    - $B(\theta) = -E\left(\frac{\partial\Psi}{\partial\theta} \mid \theta\right)$;
    - $\Omega(\theta) = E(\Psi^2 \mid \theta) = \mathrm{Var}(\Psi \mid \theta)$.

Relation with BSL

- *Consider this specific case for the QL*:
- Let

$$\psi(s_{obs};\theta) = s_{obs} - b(\theta)$$

- then we get the BSL (Stefano Cabras, Castellanos, and Ruli (2015) and S. Cabras, Castellanos, and Ruli (2014)):

$$L_Q(\theta) = \phi\left(\frac{b(\theta) - s_{obs}}{\sigma_R(\theta)}\right),$$

- where $\sigma_R^2(\theta) = Var(S \mid \theta)$.
- Is this the BSL ?

Some references on QL

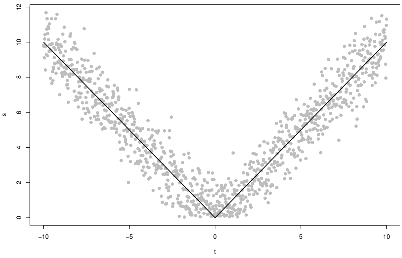
- General frequentist literature on QL with asymptotic analyses:
  - Heyde (1997),
  - Severini (2002)
- Application to Bayesian Inference: Lin (2006);
- ABC context:
  - To build a proposal for MCMC: Stefano Cabras, Castellanos, and Ruli (2015);
  - As a “synthetic likelihood”: S. Cabras, Castellanos, and Ruli (2014).

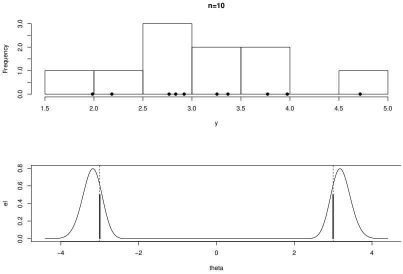
Practical implementation of the QL

- We didn't simulate m times s at each θ:
  - used a *pilot run* study with m simulations from $\pi(s(y)|\theta)$ coupled with
  - *regressions techniques* to estimate $b(\theta)$ and its derivatives (e.g. splines);
  - sometimes *just linear regressions*:
    - as $n \to \infty$ $b(\theta)$ becomes linear in a neighbourhood of the true θ_0;
  - a system of linear equations to deal with $d > 1$;
- Note that: when $b(\theta)$ and $\sigma_R(\theta)$ have an analytical approximation, also the *QL has an analytical expression*.

Examples of QL when s is not normal:

- $y \sim N(|\theta|, 1)$, with $\pi(\theta) \propto 1$,
- for $\theta \in \mathbb{R}$ and $s = |y|$
- this is the $b(\theta)$ function and the corresponding posterior:





Missspefication := there is no $\theta : b(\theta) = s_{obs}$

- In standard ABC with ϵ large, the problem is swept under the carpet;
- Maybe this is just a matter of language but:
  - why insist on a misspecified model given that we know it is misspecified?
- Author proposes different solutions for this:
  - *different estimations of the Variance/Covariance* matrix of the BSL or
  - *model embedding* by introducing an additional shrinkage parameter γ making the model more reach than the original one with θ only:
    - No wonder a more complex model fits the data better.

- (I think) that we have different interpretations of θ because:
  1. we start from an original model $\pi(y|\theta)$;
  2. define a Likelihood-based on an S, $L_{O:ABC,BSL,etc}(\theta) = \pi(s(y)|\theta)$;
  3. modify the likelihood with some $L_M(\theta)$

  - In (1)-(3), the notation θ never changes:
    - it should be $\pi(y|\theta)$, $L_O(\theta')$, $L_M(\theta^*)$;
    - By the way, for each version of the summary stat., $S^{(k)}$ it should be $\theta^{(k)}$;
    - *parameters should be interpreted within the model.*

- A result of this (I think) is in Figure 4 in Frazier and Drovandi (2021) for the MA model.

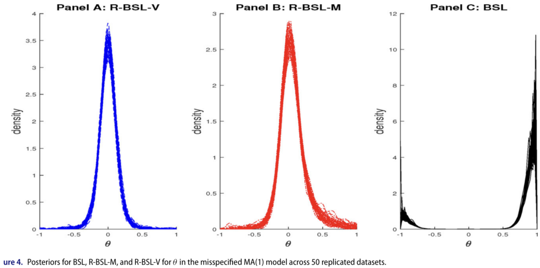


Figure 1: Different posteriors for θ

- Why are these three different posteriors (same prior) for the *supposed the same parameter* θ in the MA(1) model?
- Is this only a matter of posterior approximation?
- Is missspecification (maybe) originated by a problem of model identifiability due to non orthogonal parametrization ?

References

Cabras, S., M. E. Castellanos, and E. Ruli. 2014. "A Quasi Likelihood Approximation of Posterior Distributions for Likelihood-Intractable Complex Models." *Metron* 72 (2): 153–67.

Cabras, Stefano, Maria Eugenia Castellanos, and Erlis Ruli. 2015. "Approximate Bayesian Computation by Modelling Summary Statistics in a Quasi-Likelihood Framework." *Bayesian Analysis* 10 (2): 411–39.

Chaudhuri, Sanjay, Subhroshekhar Ghosh, David J Nott, and Kim Cuc Pham. 2020. "On a Variational Approximation Based Empirical Likelihood ABC Method." *arXiv Preprint arXiv:2011.07721*.

Frazier, David T, and Christopher Drovandi. 2021. "Robust Approximate Bayesian Inference with Synthetic Likelihood." *Journal of Computational and Graphical Statistics* 30 (4): 958–76.

Frazier, David T, Christopher Drovandi, and David J Nott. 2021. "Synthetic Likelihood in Misspecified Models: Consequences and Corrections." *arXiv Preprint arXiv:2104.03436*.

Frazier, David T, David J Nott, Christopher Drovandi, and Robert Kohn. 2022. "Bayesian Inference Using Synthetic Likelihood: Asymptotics and Adjustments." *Journal of the American Statistical Association*, no. just-accepted: 1–28.

Heyde, C. C. 1997. *Quasi-Likelihood and Its Application: A General Approach to Optimal Parameter Estimation*. Springer Verlag.

Lin, L. 2006. "Quasi Bayesian Likelihood." *Statistical Methodology* 3 (4): 444–55.

McCullagh, P. 1991. "Quasi-Likelihood and Estimating Functions." In *Statistical Theory and Modelling*, edited by D. V. Hinkley, N. Reid, and E. J. Snell, 265–86. Chapman; Hall: London.

Severini, T. A. 2002. "Modified Estimating Functions." *Biometrika* 89 (2): 333–43.