Probabilistic Models for Integration Error

Chris. J. Oates[1,5], Steven Niederer[2], Angela Lee[2], François-Xavier Briol[3], Mark Girolami[4,5]

[1]Newcastle University, [2]King's College London, [3]University of Warwick, [4]Imperial College London, [5]Alan Turing Institute

Overview

This paper studied the **numerical computation of integrals**

$$\int_\Omega f(x)\, p(\mathrm{d}x)$$

representing estimates or predictions, over the output $f(x)$ of a computational model with respect to a distribution $p(\mathrm{d}x)$ over uncertain inputs x to the model. For the functional cardiac models that motivated this work, **neither f nor p possess a closed-form** and evaluation of either requires ≈ 100 CPU hours, precluding standard numerical integration methods.

Motivation

Recall that the standard Monte Carlo confidence interval for an integral

$$\left(\bar{f} - t^* \frac{s}{\sqrt{n}}, \bar{f} + t^* \frac{s}{\sqrt{n}}\right) \qquad (1)$$

can fail in an arbitrarily dramatic manner when n is small [1]:

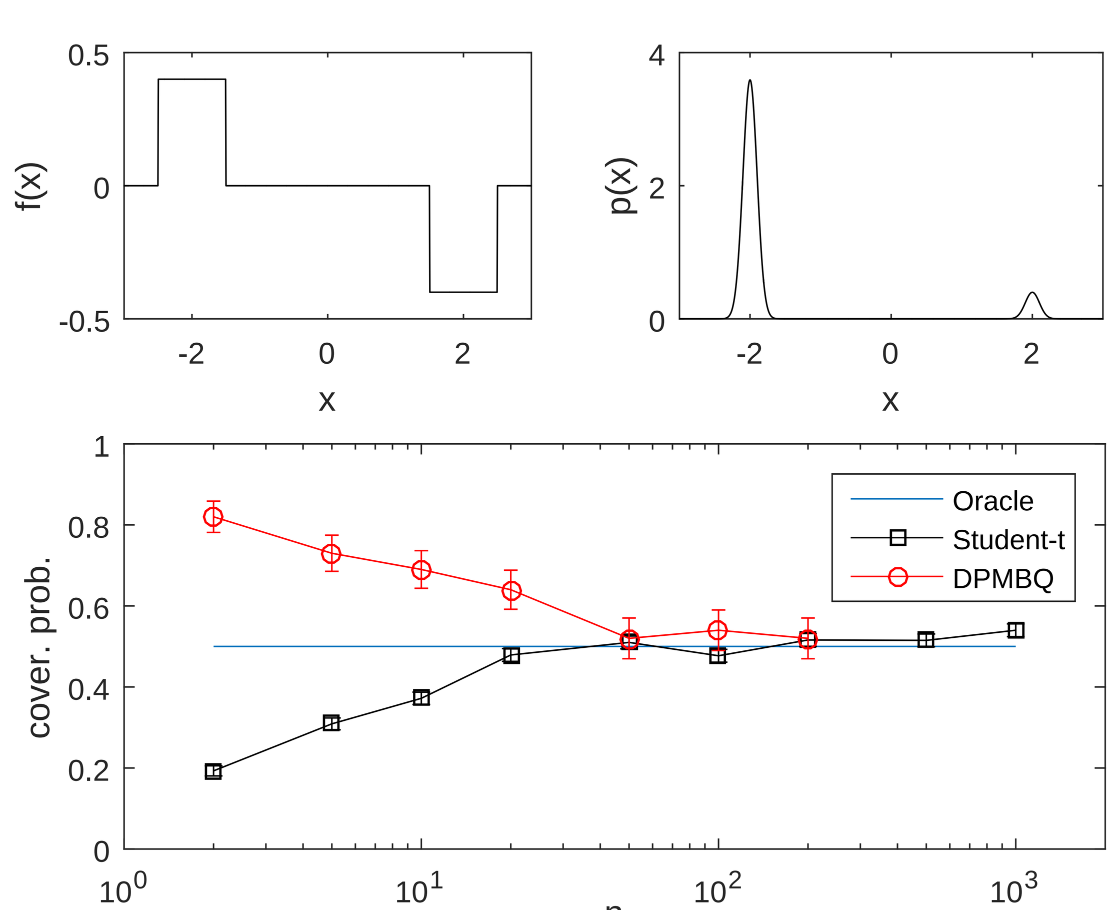

Figure 1: Consider drawing samples $\{x_i\}_{i=1}^n$ from $p(\mathrm{d}x)$ and computing an (asymptotic) 50% confidence interval (Eqn. 1) for $\int f(x)p(\mathrm{d}x)$ based on the data $\{f(x_i)\}_{i=1}^n$. The figure shows that it is trivial to construct an example for which Eqn. 1 is severely over-confident.

Building Blocks

Our aim is to use additional prior information in the construction of the confidence interval.

Gaussian Process $f \sim \mathrm{GP}(m,k)$

$$\begin{bmatrix} \vdots \\ f(x_i) \\ \vdots \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \vdots \\ m(x_i) \\ \vdots \end{bmatrix}, \begin{bmatrix} & \vdots & \\ \dots & k(x_i,x_j) & \dots \\ & \vdots & \end{bmatrix} \right)$$

Dirichlet Process $P \sim \mathrm{DP}(\alpha, P_b)$

$$[P(S_1), \dots, P(S_n)] \sim \mathrm{Dir}(\alpha P_b(S_1), \dots, \alpha P_b(S_n))$$

for a partition $\Omega = \dot\cup_{i=1}^n S_i$.

DP Mixture Model $p \sim \mathrm{DPMM}(\psi, \alpha, P_b)$

$$p(\mathrm{d}x) = \int_\Omega \psi(\mathrm{d}x; \phi) P(\mathrm{d}\phi)$$

where e.g. $\psi(x;\phi)$ is the p.d.f. for $\mathrm{N}(\phi_1, \phi_2)$.

Compatible (k, ψ) Pairs

Result from Bayesian quadrature [2]:

$$\int f \mathrm{d}p \mid p, \{(x_i, f(x_i))\}_{i=1}^n \sim \mathrm{N}(\mu_n, \sigma_n^2)$$

where (e.g.) the mean can be expressed:

$$\mu_n = \left[\dots \int k(\cdot, x_i) \mathrm{d}p \dots\right] \begin{bmatrix} & \vdots & \\ \dots & k(x_i,x_j) & \dots \\ & \vdots & \end{bmatrix}^{-1} \begin{bmatrix} \vdots \\ f(x_i) \\ \vdots \end{bmatrix}$$

Stick-breaking for DPMM [3]:

$$p(\mathrm{d}x) = \sum_{j=1}^\infty w_j \psi(\mathrm{d}x; \varphi_j)$$

where the w_j and φ_j can be sampled.

Thus required to have a closed-form for the integral:

$$\int_\Omega k(\cdot, x_i) \mathrm{d}p = \sum_{j=1}^\infty w_j \underline{\int_\Omega k(x, x_i) \psi(x; \varphi_j) \mathrm{d}x}$$

Method in a Nutshell

Take independent priors $f \sim \mathrm{GP}(m,k)$ a Gaussian process and $p \sim \mathrm{DPMM}(\psi, \alpha, P_b)$ a Dirichlet process mixture model. Then form a posterior $(f,p) | \{(x_i, f(x_i))\}_{i=1}^n$ and extract the marginal $\int f \mathrm{d}p \,| \{(x_i, f(x_i))\}_{i=1}^n$. The latter provides our model for integration error, which contracts at $O(n^{-1/4})$.

Mathematical Section

Let $\mathcal{H}$ denote the RKHS with kernel k. Suppose:

1. f belongs to $\mathcal{H}$ and k is bounded on $\Omega \times \Omega$.

Let P_0 denote the true mixing distribution. Suppose:

1. $\psi(\mathrm{d}x; \varphi) = \mathrm{N}(\mathrm{d}x; \varphi_1, \varphi_2)$.
2. P_0 has compact support $\mathrm{supp}(P_0) \subset \mathbb{R} \times (\underline{\sigma}, \overline{\sigma})$ for some fixed $\underline{\sigma}, \overline{\sigma} \in (0, \infty)$.
3. P_b has positive, continuous density on a rectangle R, s.t. $\mathrm{supp}(P_b) \subseteq R \subseteq \mathbb{R} \times [\underline{\sigma}, \overline{\sigma}]$.
4. $P_b(\{(\varphi_1, \varphi_2) : |\varphi_1| > t\}) \leq c \exp(-\gamma |t|^\delta)$ for some $\gamma, \delta > 0$ and $\forall\, t > 0$.

Denote with $\mathbb{P}_n$ the posterior marginal over $\int f \mathrm{d}p$ given $\{(x_i, f(x_i))\}_{i=1}^n$.

Then for all $\delta > 0$, $\mathbb{P}_n[(p_0(f_0) - \delta, p_0(f_0) + \delta)] = 1 - O_P(n^{-1/4+\epsilon})$ where the constant $\epsilon > 0$ can be arbitrarily small.

Application to a Cardiac Model

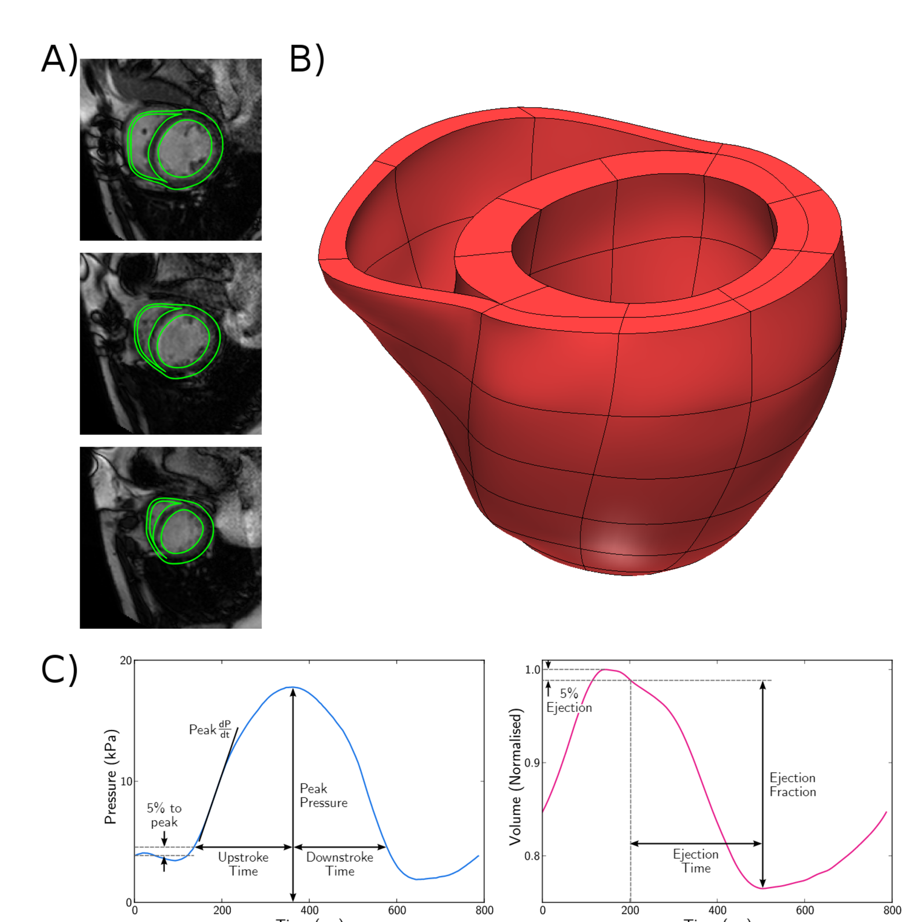

Figure 2: Computational cardiac model. This multi-scale, multi-physics model contains 10 input parameters which, for assessment purposes, must be integrated out.

Conclusion

The use of prior information can help to avoid over-confidence in numerical estimation of an integral. However, the performance of our Dirichlet process mixture Bayesian quadrature (DPMBQ) method depends critically on the appropriateness of the prior model.

References

[1] A O'Hagan.
Monte Carlo is Fundamentally Unsound.
Journal of the Royal Statistical Society, Series D, 36(2/3):247–249, 1987.

[2] A O'Hagan.
Bayes–Hermite Quadrature.
Journal of Statistical Planning and Inference, 29(3):245–260, 1991.

[3] J Sethuraman.
A Constructive Definition of Dirichlet Priors.
Statistica Sinica, 4(2):639–650, 1994.

Acknowledgements

CJO and MG were supported by the Lloyds Register Foundation Programme on Data-Centric Engineering. SN was supported by an EPSRC Intermediate Career Fellowship. FXB was supported by the EPSRC grant [EP/L016710/1]. MG was supported by the EPSRC grant [EP/J016934/1, EP/K034154/1], an EPSRC Established Career Fellowship, the EU grant [EU/259348] and a Royal Society Wolfson Research Merit Award. This material was based upon work partially supported by the National Science Foundation (NSF) under Grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute. Opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Contact Information

- Web: oates.work
- Email: chris.oates@ncl.ac.uk

Newcastle University

The Alan Turing Institute