Bayesian analysis is a statistical method that uses probabilities to measure uncertainty about unknown data in order to draw a proper inference. The analysis of complex statistical models using Bayesian inference has become more prominent in recent years. This is partly due to the increasing availability of computing power. Another factor is the discovery of a number of simulation techniques that made Bayesian inference possible for complex models. However, many of these techniques are computationally demanding. This research attempts to enhance Bayesian applications performance by using the ParInt software package.

Bayesian models analysis often requires the evaluation of complicated multidimensional integrals. For higher-dimensional non-linear models, the only practical methods for analysis are based on stochastic simulation techniques such as Monte Carlo (MC), Quasi-Monte Carlo (QMC) and Markov chain Monte Carlo (MCMC) techniques. These are known to be computationally intensive, with some analyses requiring days of CPU time on powerful computers.

Methods

The purpose of ParInt is to approximate multidimensional integrals of the form:

\[ \int_\Omega f(x) \, dx \]

The output contains an approximation \( \hat{Q} \) and an absolute error estimate \( \|f\|_{\infty} \) satisfying:

\[ |Q - \hat{Q}| \leq \|f\|_{\infty} \cdot E \]

where \( E \) is the absolute and the relative tolerance error, respectively.

ParInt applies an adaptive region subdivision algorithm (Figure 2) that divides the domain \( \Omega \) into subregions (Figure 3) for evaluation.

Results

We selected several applications [4] from various areas to be implemented in C and evaluated using ParInt:

- Medical
- Climate
- Economics

We consider two examples from Evans and Swartz [5,6] to illustrate our results.

Example 1: Linear model

This model represents simulated data and is specified as:

\[ y = \theta x + \epsilon \]

where \( y \) is the response variable, \( \theta \) is the unknown parameter, \( x \) is a predictor variable, and \( \epsilon \) is the error term.

Example 2: Contingency table

This example analyzes a cross-classification of 132 long-term schizophrenic patients in a table with three rows and three columns. The results are accurate compared to the exact results. Observe that the absolute error estimates drop considerably with the higher number of samples used (Figure 7).

Conclusion & Future work

- Bayesian models require solving high-dimensional multivariate integration problems.
- We implemented a number of Bayesian problems using ParInt.
- ParInt computations deliver accurate and efficient results. Running the applications on ParInt reduces the execution time considerably.
- We plan on adding tools for enhancing Bayesian inference solutions in ParInt.
- We will continue to investigate the role of high Performance Computing in analyzing complicated Bayesian models.

References