

# INLA for Performance Ranking Algorithms

Tristan Sones-Dykes

Date: November 18, 2023

## Theoretical Background

### Bayesian inference

In Bayesian statistics, all unknown quantities of a model are treated as random variables. The overall aim is to compute the joint posterior distribution of these parameters  $\theta$  given the observed data  $X = \{x_1, \dots, x_n\}$ .

This relies on Bayes Theorem:

$$\pi(\theta | X) = \frac{\pi(X | \theta)\pi(\theta)}{\pi(X)}$$

Where  $\pi(X | \theta)$  is the likelihood of the data given the parameters,  $\pi(\theta)$  is the prior distribution of the parameters and  $\pi(X)$  is the marginal likelihood, acting mostly as a normalising constant.

Summary statistics for  $\theta_i \in \theta$  can be obtained using the joint posterior distribution for  $\theta$ ,  $\pi(\theta | X)$ , as well as from the posterior marginal distribution of  $\theta_i$ ,  $\pi(\theta_i | X)$ ; this can be obtained by integrating the other parameters out of the posterior joint distribution.

### Conjugate analysis

Closed forms for these posterior distributions are only available for some models. For example, if the prior and likelihood are both Gaussian, then the posterior distribution will also be Gaussian.

This leaves a large number of models for which the posterior is not available in closed form. In these cases, we must turn to numerical methods to approximate this.

### Numerical methods

Numerical methods in general are used to estimate the various integrals in Bayesian models, for example the posterior mean of a parameter  $\theta_i$  can be estimated by:

$$\mathbb{E}(\theta_i | X) = \int \theta_i \pi(\theta_i | X) d\theta_i$$

Where  $\pi(\theta_i | X)$  is the posterior marginal distribution of  $\theta_i$ .

The most common algorithm is Markov Chain Monte Carlo (MCMC), this attempts to converge to the joint posterior distribution in order to obtain samples from it; these can then be used to estimate summary statistics. It is a powerful technique, but can be slow to converge and has performance issues with high dimensional models and hierarchical models. This is where Integrated Nested Laplace Approximation (INLA) comes in.

## Integrated Nested Laplace Approximation

INLA is a deterministic method for directly estimating the posterior marginal distribution of a parameter  $\pi(\theta_i | X)$ , rather than the joint posterior distribution  $\pi(\theta | X)$ , often doing this faster and with more accuracy than MCMC.

This comes at the cost of only being able to use INLA for Latent Gaussian Models (LGMs), which turns out to be a large class of commonly used models.

### Latent Gaussian Models

This is a class of models that encapsulates a range of models including additive models, generalised linear models (GLMs), generalised additive models (GAMs), spatial models and survival models. For our use-case we will be focusing on the GLMs, as our performance ranking algorithm - the Bradley-Terry model - is a GLM.

## Algorithmic Implementation and Complexity analysis

## Application to the Bradley-Terry Model