

# **An Algorithm for Approximating Integrated Likelihood Functions with Applications in Meta-Analysis**

**A Ph.D. Dissertation Prospectus**

Timothy Ruel

5/23/23

## **Introduction**

The research for my dissertation involves developing a novel algorithm for numerically integrating the likelihood function of a statistical model with respect to a nuisance parameter. This prospectus aims to demonstrate how the algorithm works and explain the appeal of using an integrated likelihood function over other types of pseudolikelihood functions to make inferences about the parameter of interest in a model.

The motivation behind my research developed out of the observation that the expression for the integral of a likelihood function follows a form similar to that of the calculation of the marginalizing constant of a posterior distribution.

## **Background**

Consider a sample  $\mathbf{x} = (x_1, \dots, x_n)$  drawn from a population. What can we say about the population based on  $\mathbf{x}$ ? Does it have a central point around which its values are dispersed? Are they clustered tightly around it or are they more diffuse? Questions like these are the motivation behind much of statistical inference. It becomes necessary to sacrifice a small amount of accuracy for (hopefully) a large reduction in complexity by imposing additional

assumptions on the problem. In the aggregate, these assumptions form the basis of what is known as a statistical model. The exact assumptions vary but

### What is a Likelihood Function?

Consider a model

### What is a Pseudolikelihood Function?

Nuisance parameters in a statistical model are a serious hindrance to making inferences about the parameter of interest. Consequently, statisticians have dedicated much time to developing techniques that will eliminate them from their models.

If we let

$$\Theta(\psi) = \{\theta \in \Theta \mid \varphi(\theta) = \psi\},$$

then corresponding to  $\psi \in \Psi$  is the set of likelihoods

$$\mathcal{L}_\psi = \{L(\theta) \mid \theta \in \Theta(\psi)\}.$$

Statisticians will often attempt to find a summary of the values in  $\mathcal{L}_\psi$  that does not depend on  $\lambda$ . This summary is called a **pseudolikelihood function** and can take several different forms: - Maximizing - Conditioning - Marginalizing\* - **Integrating**

**Profile likelihood**

**Conditional Likelihood**

**Marginal Likelihood**

**Integrated Likelihood**

**Why Use an Integrated Likelihood?**

The appeal of the integrated likelihood function as a means of eliminating nuisance parameters from the model is that it incorporates

**Approximating the Integrated Likelihood**

**The Zero-Score Expectation Parameter**

**Markov Chain Monte Carlo**

**The IL Algorithm**

**Applications**

**Multinomial Distribution**

**Standardized Mean Difference**

**References**