

What is the price of approximation? Saddlepoint methods in abundance estimation

There is a class of capture-recapture models where capture histories are subject to a known but non-reversible corruption. Examples of such models are the two-source model, where animals are captured by two protocols (say DNA and photo) that can be matched within protocol but not between protocols, model $M(t, \alpha)$ for misidentification, among others.

The complex nature of the likelihood functions for such models has recently led to the application of saddlepoint approximations.

Interpretation of the saddlepoint approximation as a likelihood function offers an alternative to classical maximum likelihood estimation (MLE) in situations where the true likelihood function is intractable, but the moment generating function of the underlying random variable is readily available. This is the case for such capture-recapture models. So how many animals do we sacrifice/invent when we apply the saddlepoint approximation in abundance estimation?

This study focuses on the validation of parameter estimates resulting from saddlepoint approximation. We propose a formula that evaluates the correctness of these estimates based on gradient computation of the neglected term in the first-order saddlepoint approximation. We perform confirmatory tests on the formula using capture-recapture models of interest and finally showcase the structure of an R package for estimation and validation based on saddlepoint methods.