Bayesian Adjustment for Confounding (BAC)

R package

Georgia Papadogeorgou

July 11, 2018

This document presents the full MCMC scheme for this implementation of BAC (1).

Let be the exposure of interest, the outcome of interest, and be the observed covariates. The method adopts two models: 1) a model for the exposure given all covariates, and 2) a model for the response given exposure and covariates

Assuming a sample of independent observations, the full data likelihood is equal to

**The priors assumed are the following:**

* Inverse gamma priors on residual variances:
* The BAC prior on inclusion indicators specifying that for each
* Normal priors on intercepts and coefficient of exposure in the outcome model ():
* The prior distribution on the regression coefficients of the covariates is dependent on the inclusion indicator, and is either a point mass at zero, or a normal distribution depending on whether the corresponding inclusion indicator is zero or one.

where represents a point-mass distribution at 0.

**The MCMC proceeds in the following order:**

*Sampling the residual variances*

The full conditional posterior distribution of is where

The full conditional posterior distribution of is similar and therefore omitted.

*Sampling the inclusion indicators and coefficients of covariates*

I use the following trick the avoid Metropolis-Hastings for the inclusion indicators. Let denote all parameters but . Then,

The numerator is the product of two prior probabilities, and the denominator is the posterior density of evaluated at zero (this is equal to 1 if , and a normal density if as seen below).

Based on the above, I calculate the posterior probability that is and sample from a Bernoulli.

If is sampled equal to , then the sample for in this iteration is also . Otherwise, is sampled from ) where

Similarly for the inclusion indicators and coefficients of the covariates in the exposure model.

*Sampling the intercepts*

The full conditional posterior distribution of is where

Similarly for the intercept of the exposure model.

*Sampling the coefficient of the exposure in the outcome model*

The full conditional posterior distribution of is where