**Variational Inference in Bayesian Variable Selection in Regression (BVSR)**

- A way to better quantify support for individual associations using two key parts:
    - 1 - Variational methods – reframe computation of posterior probabilities as an optimization problem by optimizing some criterion to find the distribution in a class that best matches the posterior
    - 2 - Importance sampling – to compute low-dimensional posterior of the hyperparameters, replace importance weights with a lower bound calculated using approximation from part 1
        - Note – look more into importance sampling
        - Using this means no assumption needed that hyperparameters are independent
- BVSR set-up essentially the same as in previous Stephens/Guan paper
- Decompose posterior inclusion probabilities: $PIP(k) = \int p(\gamma_k = 1|\boldsymbol{X}, y, \theta)p(\theta|\boldsymbol{X}, y)d\theta$
- The algorithm essentially consists of two nested loops:
    - Outer loop – approximate $p(\gamma_k = 1|\boldsymbol{X}, y, \theta)$ by minimizing K-L divergence
    - Inner loop – estimate $p(\theta|\boldsymbol{X}, y)$ w/ importance sampling, using results from outer loop
- Since there are no assumptions about the posterior distribution of hyperparameters, can obtain posterior correlations between hyperparameters
- In two simulations of about 1000 SNPs, one an "ideal" case with little multicollinearity, and one in a targeted region with strong collinearity, this outcomes of this method were very close to those from MCMC
- Implemented in R package "susieR"

**Mean Field Variational Bayes for High-Dimensional Regression, Sparse Priors**

- Recall – variational Bayes optimizes within a family of distributions (*variational family*) to find the one closest to the posterior
    - *Mean-field family* – a family of distributions under which the model parameters are independent
- This paper – uses a mean field family of distributions that independently assign each element of the parameter vector $\boldsymbol{\beta}$ an independent mixture of a Normal and Dirac mass at zero → mirrors form of spike-and-slab prior (but does not resemble the spike-and-slab posterior)
- Uses Laplace slabs for the prior instead of Gaussian slabs
- Algorithm proceeds by sequentially updating the model parameters by minimizing K-L divergence, iterating until convergence
    - Algorithm includes a *prioritized updating order*, where, in a pre-loop step, one computes an initial estimate of the mean vector of the variational family, and then update the parameters in decreasing order of absolute value of each coefficient's estimate
        - This is done because these algorithms can often be somewhat sensitive to update order
- Implemented in R package "sparsevb"