* **Dirichlet Process Regression**
  + Assumes that effect sizes follow a normal prior with variance (not sure if the variance is common for all coefficients or not?), i.e.
  + Hyperparameter is placed under a **nonparametric** prior using the Dirichlet process
    - * is the base distribution (Xiang’s paper assumes Gamma), is the concentration parameter that describes how G deviates from H and is inferred from the data
  + The prior can approximate well t and point-t distribution priors, a uniform step mixture, and a distribution of marginal effects (estimated from a real data set, unsure if would replicate in others but see no reason why not)
  + Simulated well with a data set of ~ 42,000 SNPs
* **Horseshoe Regularization Penalty**
  + Looks at optimization problem as finding posterior *mode* under a given prior
  + Enables local linear approximation optimization algorithms, and MCMC for posterior simulation
  + Using the posterior mode of the horseshoe yields an estimator that is sparse and nearly unbiased for large signals, but is not continuous in the data
  + They propose a proper prior that mimics the horseshoe (pole at origin, polynomial/heavy tails), calling it the **horseshoe-like prior**
    - Density is , where
  + Chart, histogram

    Description automatically generatedSince the prior is now proper, can use expectation-maximization algorithm or other local maximization techniques
  + Right – horseshoe-like prior density with
  + Can derive the distribution of the parameter as , where
  + Can also be expressed as a scale mixture of Cauchy distributions
  + Paper includes methods with the horseshoe-like prior for *EM subset selection in high-dimensional regression* as well as MCMC
  + Largest data set applied to in the paper is a real gene expression set with 3,051 genes, 38 samples
    - 🡪 could run our own simulations to see how well this method scales to larger genetic data sets
* **Variable Selection with SuSiE**
  + Uses the technique of *sum of single effects* regression
    - The *single effects* model assumes that exactly one of the predictors has a nonzero coefficient
    - Single-effect model: vector has exactly one nonzero element, where ,
  + **Sum of single effects model**:
    - Introduce L single-effect vectors , where ,
    - Then gives the sum
    - gives standard BVSR
  + Approximately equivalent to BVSR, but with two computational advantages
    - Deterministic algorithm for approximate posterior distributions
    - Simple calculation of credible sets
      * *Credible set* of level is a subset of variables with probability or greater of containing at least one effect variable (useful for quantifying uncertainty in which variables to select)
  + Given , estimating reduces to fitting a single effects model
    - Leads to an iterative algorithm the estimates with a single effects model given the current estimates for other ,
    - Effect of predictor j is
  + Model is robust to large L
    - If L is overstated, the uncertainty in the model spreads out extra effects among the covariates and makes little difference overall
    - Key signals remain the same