May 30, 2019

1 Summary

Random effects to have a nonparametric prior distribution with Dirichlet process prior.

2 Introduction

Individual i with n_i repeated measurements.

$$y_i = X_i\beta + Z_ib_i + e_i \qquad i, \dots n$$

- y_i is $n_i \times 1$
- X_i is $n_i \times p$ matrix of fixed covariates
- β is $p \times 1$ regression coefficients parameter vector
- Z_i is $n_i \times v$ covariates
- $b_i \sim N_v(0, D)$ is $v \times 1$ random effects
- $e_i \sim N_{ni}(0, \sigma^2 I_{n_i})$ is $n_i \times 1$ vector of errors
- e_i and b_i are independent

Under these assumption, assumes a distinct set of regression coefficients for each individual once the random effects are known

$$y_i|\beta, b_i \sim N_{n_i}(X_i\beta + Z_ib_i, \sigma^2 I_{n_i})$$

The distribution of random effects is usually taken to be normal. marginally

$$y_i|\beta, \sigma^2, D \sim N_{n_i}(X_i\beta, Z_iDZ_i' + \sigma^2 I_{n_i})$$

Bayesian inference for β using the marginal likelihood will depend only on(y, σ^2, D) But the nature of this dependency will be sensitive to the distributional form ascribbed to the b_i, Thus it is important to setting a good prior to b_i

3 Mixture of Dirichlet Process

$$\text{Stage 1} : x_i|\theta_i \sim D_{n_i}(h_1(\theta_i)),$$
$$\text{Stage 2} : \theta_i|\Psi_0 \sim D_w(h_2(\Psi_0))$$

- $D_s()$ is a s-dimensional multivariate distribution
- $h_1()$ and $h_2()$ are fucntions
- G general distribution, or base measure that approximates the true nonparametic shape of G
- G_0 is a w-dimensional parametric distribution
- Scalar M reflects how similar the nonparametic distribution G is to the base measureG_0

MDP model removes assumption of a parametric prior at the second stage and replaces it with a general distribution G which has Dirichlet process prior.

$$\text{Stage 1} : x_i|\theta_i \sim D_{n_i}(h_1(\theta_i)),$$
$$\text{Stage 2} : \theta_i \sim^{iid} G$$
$$\text{Stage 3} : G|M, \Psi_0 \sim DP(M \cdot G_0(h_2(\Psi_0)))$$

- As $M \to \infty$, $G \to G_0$, so that the base measure is the prior distribution of θ_i
- if $\theta_i = \theta$, the base measure is the prior distribution of θ_i too

4 Polya urn representation of Dirichlet process

- Draw of θ_1 always from the base measure
- Draw of θ_2 is equal to θ_1 with probaility p_1 and from base measure with $p_0 = 1 - p_1$
- Draw of θ_3 is equal to θ_1 with probaility p_1, equal to θ_2 with probaility p_2 and from base measure with $p_0 = 1 - p_1 - p_2$
- θ_n is equal to θ_i with probaility p_I and from from base measure with $p_0 = 1 - \sum_{i=1}^{n-1} p_i$. The value p_i's are determined from the Dirichlet process parameters.

5 Conditional distribution of θ

Conditional on the other θ's θ_i has the mixture distribution

$$p(\theta_i|x, \theta_{-i}) \propto \sum_{j \neq i} q_j \delta_{\theta_j} + M q_0 h_0(\theta_i) p(x_i|\theta_i)$$

- q_j and Mq_0 can be normailized to get the selection probabilities p_i, $i = 0, \ldots n-1$
- $p(x_i|\theta_i)$ is the smapling distribution
- δ_s is degenerate distribution with point mass at s
- g_0 is density correspond to G_0
- $q_j = p(x_i|\theta_j)$ and $q_0 = \int p(x_i|\theta) g_0(\theta) d\theta$

6 DP prior in random effect model

Model of p fixed effects and v random effets is

$$y_i|\beta, b_i, \sigma^2 \sim N_{n_i}(X_i\beta + Z_i b_i, \sigma^2 I_{n_i})$$

Letting $\tau = \sigma^{-1}$. priors are

$$\begin{aligned}
&\tau \sim Gamma\left(\frac{a_0}{2}, \frac{\lambda_0}{2}\right) \\
&\beta \sim N_p(\mu_0, \Sigma_0) \\
&b_i \sim G \\
&G \sim DP(M \cdot N_v(0, D))
\end{aligned}$$