Dependence structures for longitudinal Bayesian nonparametric models

Steve MacEachern
The Ohio State University
along with Xinyi Xu, Zhiguang Xu, and Hyoin An

ISBA 2022, Montreal

June 2022

Longitudinal data and BNP

- **Huge literature on BNP models**
  - **much concerned with conditionally IID data**

$$F \sim \textbf{BNP model}$$

$$\beta_i | F \overset{IID}{\sim} F$$

$$Y_i | \beta_i \overset{IND}{\sim} N(\beta_i, \sigma^2)$$

  - **for discrete models, "atoms" often reified**
  - **very effective as component for random effects**

$$Y_{i,j} | \beta_i, \tau_j \overset{IND}{\sim} N(\tau_j + \beta_i, \sigma^2)$$

- **Longitudinal data have more complicated dependence structures**
  - **think hard about properties we wish to have**
  - **develop strategies to construct models with desired properties**

The basics - a single distribution

- **Typical construction of a Bayesian nonparametric model**
  - **a countable collection of independent random variates**
  - **a rule for turning the variates into a distribution**
- **Usual example: Dirichlet process via Sethuraman's construction**
  - **the variates**

$$\theta_i \overset{IID}{\sim} F_0, \qquad i = 1, 2, \ldots$$
$$v_i \overset{IID}{\sim} \mathbf{Beta}(1, M), \qquad i = 1, 2, \ldots$$

  - **and the rule for constructing a distribution**

$$w_i = v_i \prod_{j<i} (1 - v_j)$$
$$F = \sum_i w_i \delta_{\theta_i}$$

- **More - Polya trees, Pitman-Yor processes, quantile pyramids**
- **Distribution is generally embedded in larger framework**

The scope of these constuctions is bigger than it seems

- **Expanding the Dirichlet process**
  - **replace Beta$(1, M)$ with zero-enriched Beta$(1, M)$**
    - **zero enrichment allows v_i to be 0 with some probability**
  - **result is still a Dirichlet process**
  - **replace Beta$(1, M)$ with one-enriched Beta$(1, M)$**
    - **one enrichment produces truncation of mixture**
  - **result may be a finite mixture model when $P(v_i = 1$ for some $i) = 1$**
  - **result may be a finite or countable mixture model when $0 < P(v_i = 1$ for some $i) < 1$**
  - **allow dependence in sequence of variates**
    - **needed to recover conventional priors for finite mixture**

Moving from one distribution to many

- **Similar construction, replace variate with stochastic process**
  - **index set is time, space, covariate, or all these**

$$\theta_{i,t} \overset{IID}{\sim} F_{0,t}, \qquad i = 1, 2, \ldots$$
$$v_{i,t} \overset{IID}{\sim} \mathbf{Beta}(1, M_t), \qquad i = 1, 2, \ldots$$

  - **and the rule for constructing a distribution**

$$w_{i,t} = v_{i,t} \prod_{j<i} (1 - v_{j,t})$$
$$F_t = \sum_i w_{i,t} \delta_{\theta_{i,t}}$$

- **The only concern is measurability**
  - **typical constructions use continuous functions of variates**
  - **also work for variates indexed by** t
- **And also works for finite set** $t_1, \ldots, t_k$

Creation of the model (DDPs)

- **Canonical construction**
  - **Gaussian process to uniform (PIT) to desired distn for variate**

$$Z_t \sim GP(\mu, K), \qquad \textbf{standard normal marginals}$$

$$\theta_t = F_{0,t}^{-1}(\Phi(Z_t)) \qquad \textbf{or} \qquad v_t = B_{1,M_t}^{-1}(\Phi(Z_t))$$

  - **example of a “copula process”**

- **Classic distribution theory generates much, much more**
  - **chi-square from GP; gamma from GP; beta from two gammas**
  - **t from normal and gamma**
  - **produces desired distns for variates, dependence differs**

- **Shared set constructions - tend to lead to discontinuity in F_t**
  - **sliding window captures portion of gamma process; normalize**
  - **increasing window**

Longitudinal data

- **A few properties**
  - **ability to refine time to arbitrarily fine scale (process in t)**
  - **nonparametric form at a given t**
  - **large (full) support at t_k given distn's at $t_1, \ldots, t_{k-1}$**
  - **continuity of distributions in t**
- **Connection of observations across t**
  - **a discrete mixture**
    - **old work on growth curves (single p models)**
    - **subject is one of several types, type determines path**
    - **path plus measurement error**
  - **continuous "movement" through distributions**
    - **time series models**
- **Both cases rely on pathwise behavior of distns**
  - **and also on the path through them for each subject**

Continuity of paths

- **Return to a Cauchy copula process (in t)**
  - **GP_1 with standard normal margins, conventional covariance**
  - **GP_2 with standard normal margins, conventional covariance**

$$T_t = \frac{Z_{1,t}}{\sqrt{Z_{2,t}^2}} = \frac{Z_{1,t}}{|Z_{2,t}|}$$

  - **T_t follows a Cauchy distn at each t**
    - **covariances determine dependence across t**
  - **For any finite set $t_1, \ldots, t_k$, the probability of dividing by 0 is 0**

- **But for an interval $[t_1, t_2]$, the probability of dividing by 0 is positive**
  - **division by 0 corresponds to $T_t = \pm\infty$**
  - **the path of T_t is discontinuous where this happens**
- **Care is needed to ensure continuity of paths**

The path of a subject

- **To ensure that $Y_{i,t} \sim F_t$, need a process with specified marginals**
  - **Unif$(0,1)$ marginal (or N$(0,1)$ marginal)**

- **Simple random subject effect version**
  - **constant percentile of distn, no variation in t**
  - **subject effect is Unif$(0,1)$ (or N$(0,1)$)**

- **Gaussian process subject effect, $Z_t \sim$ GP$(0, K)$**
  - **subject effect follows $\Phi(Z_t)$ (or Z_t)**
  - **stabilty / instability follows from K**

- **Other copula process subject effect, $T_t \sim$ CP$(\cdot)$**
  - **subject effect follows $G(T_t)$ (or $\Phi^{-1}(G(T_t))$)**
  - **non-Gaussian copula allows control over dependence structure**

- **Continuity of distns coupled with continuity of subject effect yields continuity of subject path**

Recap

- Build collection of distributions in style of DDP
  - replace variate with path of stochastic process
  - consider non-canonical constructions of processes
  - care with pathwise properties for longitudinal data
  - distns inherit properties from properties of stochastic processes
- Choose version of models
  - mixture version, along lines of finite mixture
  - percentile version, with changing percentiles
  - normal version of latter allows normal set point, reversion toward set point (usual distn theory for normals)
- Supplement model with additional structure
  - covariates, changes in mean, changes in scale
- Similarity of thought to work on GLMMs with Peter Craigmile & Jeff Gory